

REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-09-0228

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, gathering existing data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not have a valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 02/28/2009		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 03/01/2007 - 02/28/2009	
4. TITLE AND SUBTITLE Human Resource Scheduling in Performing a Sequence				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER FA9550-07-1-0356	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Roger W. Remington Shu-Chieh Wu Harold Pashler				5d. PROJECT NUMBER 2007001102	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Queensland, St Lucia, QLD 4072 Australia San Jose State University, One Washington Square, San José, California USA, 95192				8. PERFORMING ORGANIZATION REPORT NUMBER 2007001102	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research 875 North Randolph Street Suite 325, Rm 3112 Arlington, VA 22203				10. SPONSOR/MONITOR'S ACRONYM(S) AFOSR	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Behavior in military domains typically requires a sequence of decisions and actions. Yet, characteristics and limitation of cognitive processing are typically based on discrete-trial laboratory studies. The broad objective of this work was to bridge the gap between basic science and applications by: (1) exploring the suitability of central bottleneck models as the basis for computational models of sequence behavior, (2) identifying emergent properties in scheduling behavioral sequences; and (3) determining if and how sequence execution is optimized. Studies examined eye movements and manual responses to sequences of speeded choice response time tasks arrayed linearly on a visual display. A consistent emergent property was discovered in the deferral of the first response, which was shown to be a strategy only loosely linked to resource constraints. Given this strategy central bottleneck theory provided an accurate account of sequence execution. Deferring the first response may represent an optimal response to stochastic fluctuations in the duration of internal processes. The distribution of eye fixation was well fit by reinforcement learning models, evidence for optimality with respect to target probability.					
15. SUBJECT TERMS central bottleneck; cognitive modeling, saccadic eye movements, reinforcement learning, optimal models					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Roger W. Remington
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code) +61733656412

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

15 JUL 2009

DTIC Data

Page 1 of 2

Purchase Request Number: FQ8671-0800279

BPN: F1ATA07242B039

Proposal Number: 07-NL-112

Research Title: HUMAN RESOURCE SCHEDULING IN PERFORMING A SEQUENCE OF DISCRETE RESPONSES

Type Submission: Final

Inst. Control Number: FA9550-07-1-0356P00002

Institution: UNIV OF QUEENSLAND (AUSTRALIA)

Primary Investigator: Prof Roger Reminton

Invention Ind: none

Project/Task: 2313B / X

Program Manager: Jun Zhang

Objective:

The goal is to develop a theory of behavior composition that describes how sequences of behaviors are constructed from component cognitive, perceptual, motor operations. Efficient scheduling of human cognitive resources is an important concern for skilled performance. Data from commonly used single discrete trial paradigms cannot be generalized to complex tasks that require a coordinated schedule of actions. Through experiments and computational model analyses the investigators will determine whether behavioral sequences are simply an iteration of the single stage operations, or whether there are emergent properties, and finally whether human resource allocation for scheduling is optimally planned. Theory and experimental data on this issue is lacking, and the proposed work will break new empirical and theoretical ground on this important problem.

Approach:

This research will employ behavioral experiments and computational modeling of the experimental results to test models of composition for sequences of actions. A series of six experiments will be conducted to examine preparation effects on response time, demands on scheduling of eye movements, time pressure, and effort minimization. Tasks that require concurrent use of multiple resources will be examined (reading and typing are real world examples). A theory of human behavior composition must address the resources to be scheduled, the constraints on scheduling, and the strategies that govern scheduling. A computational modeling technique that automatically constructs behavior sequences by scheduling primitive cognitive operations will be examined, called the CMP-GOMs model. Performance produced by principles of optimal scheduling that maximizes resources will be compared with human performance.

Progress:

Year: 2007 **Month:** 02

Not required at this time.

Year: 2008 **Month:** 03

This progress report covers Year Two, whose principal goals were: (1) identify emergent properties of sequence execution, (2) examine optimality in visual sampling, (3) extend, refine, or rethink the just-in-time scheduling model developed in Year One. One of the most robust and large emergent effects is the elevation of response time to the first stimulus (RT1) seen in sequences. In our experiments RT1 is typically over 600 ms slower than the subsequent inter-response time (IRI) intervals, and can be as much as double the RT for a single response to a single stimulus. The principal questions are whether this elevation represents some emergent cost of programming a sequence, a reflection of the inefficient processing of the first item in a sequence, or reflects a strategic use of internal resources by the participant.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

15 JUL 2009

DTIC Data

Page 2 of 2

Progress:

Year: 2008 **Month:** 03

Several experiments have converged on an explanation of RT1 as a strategic effect, specifically one of buffering one stimulus until the next has been identified. Optimality of eye fixation patterns in visual search was examined in an experiment based on the reinforcement learning paradigm, yielding good fits to softMax transformation of location probability. Computational modeling has focused on a revision of the previous computational model based on just-in-time assumptions. This earlier model has been significantly altered as a result of the empirical evidence collected this year. A new central bottleneck model is being developed which will take into account strategic aspects of sequence execution and emphasizes the timing relationship between motor responses and eye fixations. Cognitive architectural assumptions other than central bottleneck are also being examined in light of this year's empirical findings. Three papers are currently in preparation that summarize empirical findings and resultant computational models.

Accomplishments

Our Year Two empirical investigations of emergent properties in sequence execution focused on two issues: sources of RT1 elevation and the contribution of eye movement planning and execution to observed inter-response intervals (IRIs). Previous results identified a substantial start-up delay in initiating a sequence of discrete responses (referred to hereafter as RT1 elevation). The source of this elevation has important consequences for understanding the scheduling of underlying perceptual, cognitive, and motor resources. For example, RT1 could include components related to oculomotor or motor planning for the sequence; these should be sensitive to features of the subsequent stimulus list. Alternatively, RT1 could reflect slowed central processing similar to "first item" effects seen in task switching experiments. Or, it could result from a strategy to buffer one response until the next stimulus is identified (or has completed response selection). We also sought to determine whether the necessity to make a sequence of eye fixations itself carried a cost that would be reflected in RT1 or IRI.

A control experiment was conducted using gaze contingent displays to prevent perceptual processing of the next item prior to the response to the previous. Results showed RT1 still approximately 400 ms slower than its single item comparison. Several studies were done to assess the contribution of preparation. In all, they agreed with the control study in concluding that preparation for performing the sequence can account for 200-300 ms of RT1 elevation. We hypothesize that the residual RT1 effect is attributed to a strategy of buffering the first response (or two) and completing response selection on S2 prior to responding to S1. This buffering could serve two purposes. First, it could assure that responses can be made with reference to a local buffer without waiting for new information to be input. This may allow low-level mechanisms to control response execution resulting in a regular periodic sequence of responses. Second, buffering could allow the system to adjust to the processing dema

Year: 2009 **Month:** 06 **Final**

2007-2008: Identified emergent properties in sequence execution; modeled strategic goal of minimizing delays in response selection; demonstrated the suitability of central bottleneck models in accounting for sequence behavior. 2008-2009: Demonstrated that response selection is completed prior to eye movement initiation, resolving long-standing theoretical conflict; developed quantitative bottleneck models of sequence execution; demonstrated success using reinforcement learning models to capture allocation of attention in visual search

Final Report for AFOSR Grant Agreement #: FA9550-07-1-0356

Period of Report: 1 March 2007 – 28 February 2009

AFOSR Scientific Officer: Dr. Jun Zhang

Title:

*Human Resource Scheduling in Performing a Sequence of
Discrete Responses*

Principle Investigators:

Roger Remington
University of Queensland
School of Psychology
St Lucia, QLD 4072 AUSTRALIA

Shu-Chieh Wu
San Jose State University and NASA Ames Research Center

Harold Pashler
University of California, San Diego

20090723669

Objectives

Behavior in military domains typically requires a sequence of decisions and actions. Yet, characteristics and limitation of cognitive processing are typically based on discrete-trial laboratory studies. The broad objective of this work was to bridge the gap between basic science and applications by: (1) exploring the suitability of central bottleneck models as the basis for computational models of sequence behavior, (2) identifying emergent properties in scheduling behavioral sequences; and (3) determining if and how sequence execution is optimized. Studies examined eye movements and manual responses to sequences of speeded choice response time tasks arrayed linearly on a visual display. A consistent emergent property was discovered in the deferral of the first response, which was shown to be a strategy only loosely linked to resource constraints. Given this strategy central bottleneck theory provided an accurate account of sequence execution. Deferring the first response may represent an optimal response to stochastic fluctuations in the duration of internal processes. The distribution of eye fixation was well fit by reinforcement learning models, evidence for optimality with respect to target probability.

Status of Effort

The project met all original goals of identifying strategic emergent properties in sequence execution and identifying where sequence execution can be considered optimal. In addition, we have demonstrated that the pattern of eye fixation is optimal in that it is an adaptive response to contingencies, well fit by reinforcement learning models. A new central bottleneck model was developed which will take into account strategic aspects of sequence execution and emphasizes the timing relationship between motor responses and eye fixations. Two papers (attached) are nearly completed and will be submitted shortly.

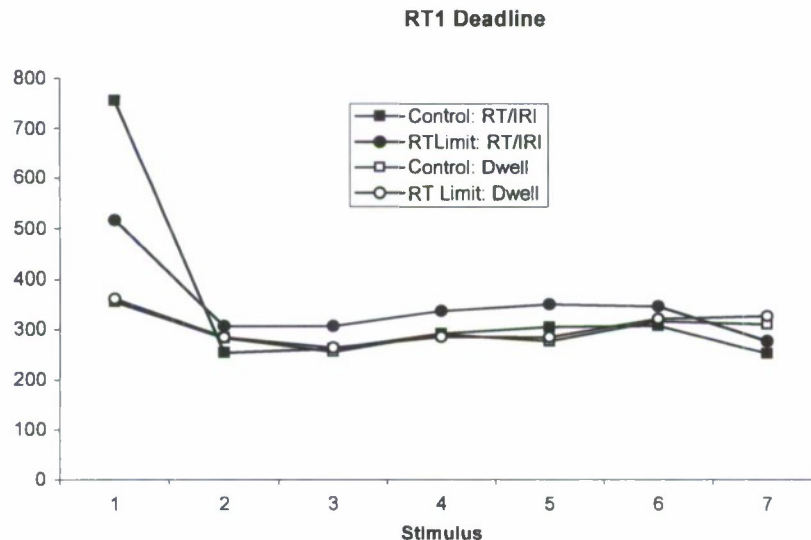
Accomplishments

The two attached drafts of papers soon to be submitted describe findings regarding the presence of one key emergent property, elevation of the first response in a sequence, along with computational models of possible resource scheduling of the eyes and hands. Briefly, the elevation of the first response is a strategy that is not dictated primarily by resource conflicts. Instead, it seems to reflect a tendency to make a regular sequence of manual responses closely coupled to the timing of saccadic eye movements. The timing of saccades is well fit by assuming that most if not all of central processing is completed prior to moving the eyes. It is not yet clear whether central processing must be completed prior to moving the eyes, or whether this reflects a strategic choice to separate the processing of adjacent items to avoid interference.

Strategic Deferral

Other critical features of sequence processing have been investigated whose experiments are in the early stages of write up. In one experiment we pursued the issue of strategic elevation of RT1 by imposing a deadline on the response time for RT1. Subjects are informed that they must respond to the first item in less than 600 ms, which is approximately the time taken to respond to only the first item in

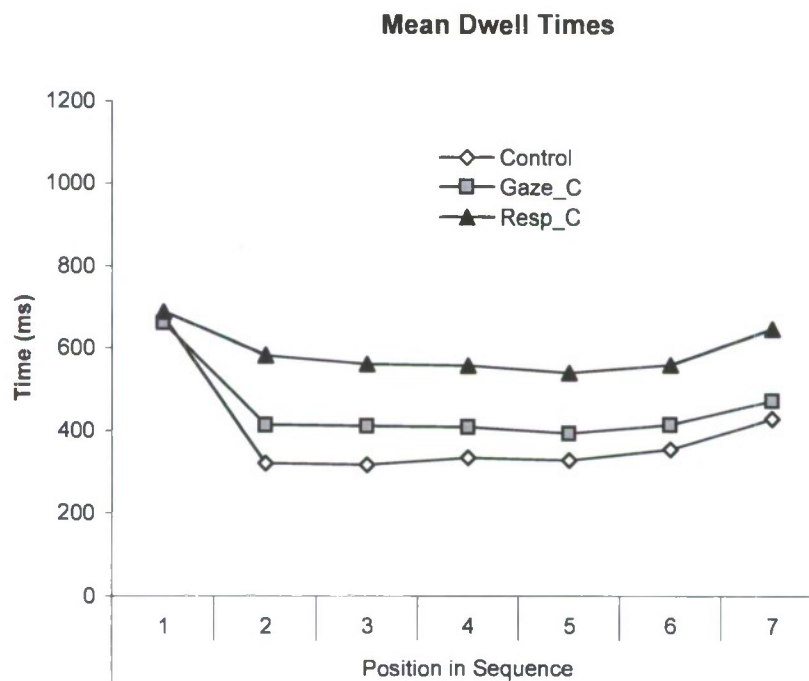
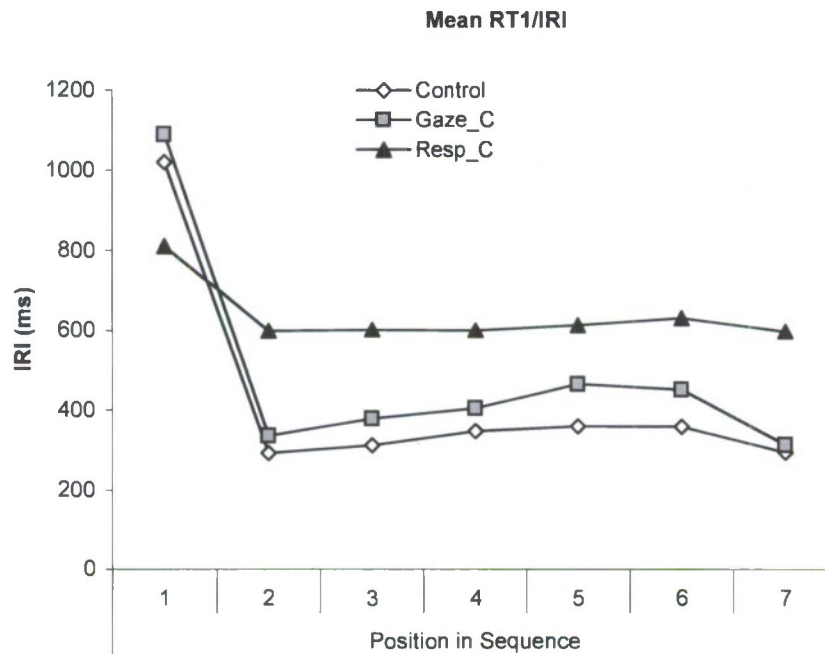
previously reported control trials (see attachment 1). In pilot studies, subjects found it very difficult to match the deadline. Two subjects were extensively trained with the deadline procedure, eventually being able to produce over 95% of first responses under the deadline. They were then tested for one session with the deadline, followed by one session without the deadline. The results below show the mean data for the two trained subjects. Filled symbols are the RT1 and mean IRI for each stimulus; the open symbols the mean dwell. There was no effect of the RT limit on Dwell. When the limit was in force RT1 was reduced from almost 800 ms to approximately 550 ms, nearly equivalent to the 575 ms observed earlier (see attachment 1 Experiment 5) for a simple response to the first item without doing anything with the remaining sequence. Mean IRI shows a small increase with deadline for all position. This may reflect a different strategy when fast responses to S1 are called for, one that rearranges the preferences for moving the eyes relative to processing the stimulus. Nonetheless, when responding quickly to S1 there was no evidence of a significant loss of efficiency, or a dramatic disruption of performance expected if RT1 elevation was truly a consequence of resource conflicts.



To determine if the small elevation in IRI for the RT Limit condition represented a significant departure from how a normal preview condition would be executed, we ran an additional experiment comparing three conditions: 1) a standard Preview condition where all stimuli were visible at trial outset; 2) a Gaze Contingent condition in which each stimulus position contained a non-informative placeholder that would change to a target stimulus once fixated; and 3) a Response Contingent condition in which the placeholder for the next stimulus would disappear to reveal the target only after the response to the previous stimulus had been made. Three of the 6 subjects performed the Preview condition followed by the Gaze Contingent, while the other three performed the Preview followed by the Response Contingent condition. A comparison of the standard Preview to the Gaze Contingent was a check to see how much information may have been processed prior to the eye movement despite the design features, which were intended to limit that. A comparison of the Response Contingent with the Preview and Gaze Contingent directly examines the benefit of

overlap and provides a comparison of the deadline to see whether the RT1 deadline eliminated some or all of the preview benefit.

The following two figures show the mean RT1/IRI and mean Dwell times, respectively, for each of the three conditions.



For statistical analyses, the Preview condition was tested against the experimental condition for each subject using a paired t-test. For manual responses, mean IRI was relatively constant across stimulus position so IRI for each subject was averaged for stimuli S2-S6. These positions were chosen to avoid issues with RT1 deferral as well as last item effects, which we reported earlier (see attachment 1). RT1 was over 200 ms faster in the Response Contingent than the experimental conditions, yet with so few subjects comparisons of the experimental to control condition failed to reach significance either for the Gaze Contingent ($t = 1.4$, $df = 2$, $p \leq .29$) or for the Response Contingent conditions ($t = 1.5$, $df = 2$, $p \leq .27$). However, there was a significant elevation of mean IRI in the Response Contingent compared to the Preview condition ($t = 32.39$, $df = 2$, $p < .0001$ 2-tailed). Mean IRI for positions 2 thru 6 (eliminating position 7 to avoid the "last item effect") was 331 ms in the Preview condition compared to 607 ms in the Response Contingent condition. Mean Dwell time was slightly elevated in the Gaze Contingent condition compared to the Preview ($t = 2.19$; $df = 2$, $p \leq .16$). With additional subjects this could well become significant. Mean Dwell in the Response Contingent was significantly elevated ($t = 32.73$; $df = 2$; $p < .0001$).

Comparison of the RT1 deadline results with those of the Response Contingent condition show clearly that speeding RT1 has relatively little effect on subsequent IRI or Dwell compared to a no-preview condition. At best then there is a small reduction of overlap but far from the effect seen when overlap is eliminated. Further testing is in progress to more fully flesh out these effects. If the current pattern holds, it will be additional evidence that the elevation of RT1 is not dictated by internal resource conflicts, but is an emergent strategy when executing a sequence.

Strategies in Information Acquisition

As part of an expanded effort we have begun to examine not only the possibility that resources are scheduled in optimal ways, but to understand as well the overall pattern of information acquisition. In previous reports we have described an experiment that examined optimality in information sampling. This work, and speculations on its implications are described in the three final attachments. Briefly, in the "Finding Happiness" experiment, participants scanned a visual display to uncover a target. Eye movements were monitored and the region being fixated was uncovered. Search continued until the fixated region contained the target. The amount of the reward was related to the length of time taken to find the target. Locations differed in the base probability of containing a target. The question investigated was whether people's fixations would become optimal with respect to the underlying probabilities. The results showed clearly that they did indeed. By the end of the first session the proportion of fixations on a location was almost precisely predicted by a SoftMax transformation of the underlying probability. SoftMax functions have been shown successful in reinforcement learning paradigms in accommodating the tension between exploiting areas, which have previously been shown to be high value, with exploring new regions. This is the first demonstration to our knowledge of a fit with eye fixations, and supports work done decades ago on optimal monitoring of cockpit displays. Further experiments on optimal search are planned and we will attempt to integrate them into the coordination results to provide a model of the coordination of

mental resources that includes decisions of whether to choose to exploit or explore on any given sample.

We have recently designed and piloted a new paradigm for examining optimal search. The central goal is to understand whether the adaptation seen in the Finding Happiness experiment reflects a conscious adaptation that affects only fixation probability, or whether this adaptation is better characterized as occurring with practice. To do this we examined how sampling behavior is affected by instructing the subjects of the likely location for a target, compared to a no-instruction condition where through practice they learned the probability associated with each location. In addition, the task was changed to force an extended decision on the observer. This was done to test whether the decision threshold for determining a target was present was also affected by the probability of a target being in that location. That is, does a likely location both increase its probability of being sampled and simultaneously decrease the information required to make a target present response? For reasons explained below we report here total dwell time, the sum of the durations of all fixations on each location.

In this pilot experiment we examined whether explicit cueing of a location led to the same type of set as learning an implicit probability. Subjects viewed a set of 4 dials space evenly around the perimeter of a circle with a small fixation cross at its center. Each circle represented a dial measuring the quantity of an unspecified substance. Running through the center of each circle was a horizontal line with tick marks at regular intervals. At the beginning of a trial a vertical line was presented at one of the tick marks on the horizontal line. When the trial commenced the lines began to be perturbed horizontally as the result of adding random gaussian noise. At some point in the trial a step function was added to one of the dials resulting in a mean displacement of the vertical line. Subjects were instructed to press a key when they detected this displacement, following which they indicated which dial was displaced. The dials were sufficiently far into the periphery that subjects had to move their eyes to determine the position of the vertical line in a dial. Eye movements were recorded using an EyeLink 1000 sampling at 240 Hz.

There were three cueing conditions. In the Explicit condition a central arrow was presented on each trial indicating which of the 4 locations was the most likely location of the target step change. On 70% of trials the target step change occurred in the cued location, and 10% in each of the other three locations. In the Implicit condition, no instructions were provided. Instead, the most likely target location was fixed for each subject for that block of 80 trials. Subjects were given no explicit instructions regarding the probability of any location. However, feedback was given after each response to indicate where the target had occurred. Thus, given the discrepancy in probabilities, it was not difficult to determine the likely location after a few trials. In the Random condition, all locations were equally likely and no instructions were given beforehand. Six subjects were tested for 240 trials, 80 in each of the three conditions. The Explicit condition was always tested first, the Random and Implicit conditions alternated between second and third.

We have conducted a preliminary analysis of the data focusing on the effects of explicit cueing, implicit learning, and cue validity on mean RT and total fixation duration. Total fixation duration is the sum of the times for all fixations on a given

location from the onset of the trial to the response. In the first graph below target response time is plotted as a function of cue validity for the three cueing conditions: Explicit (central cue), Implicit (fixed location preference), and Random. RT for the Random condition is graphed in the "Invalid" as there were no cues in that condition, and hence no valid trials.

The second graph shows total fixation duration on Valid and Invalid trials for the Explicit and Implicit conditions, as well as the predicted optimal fit under the SoftMax model. SoftMax predictions were calculated using the formula

$$p'_i = e^{p_i} / \sum e^{p_i}$$

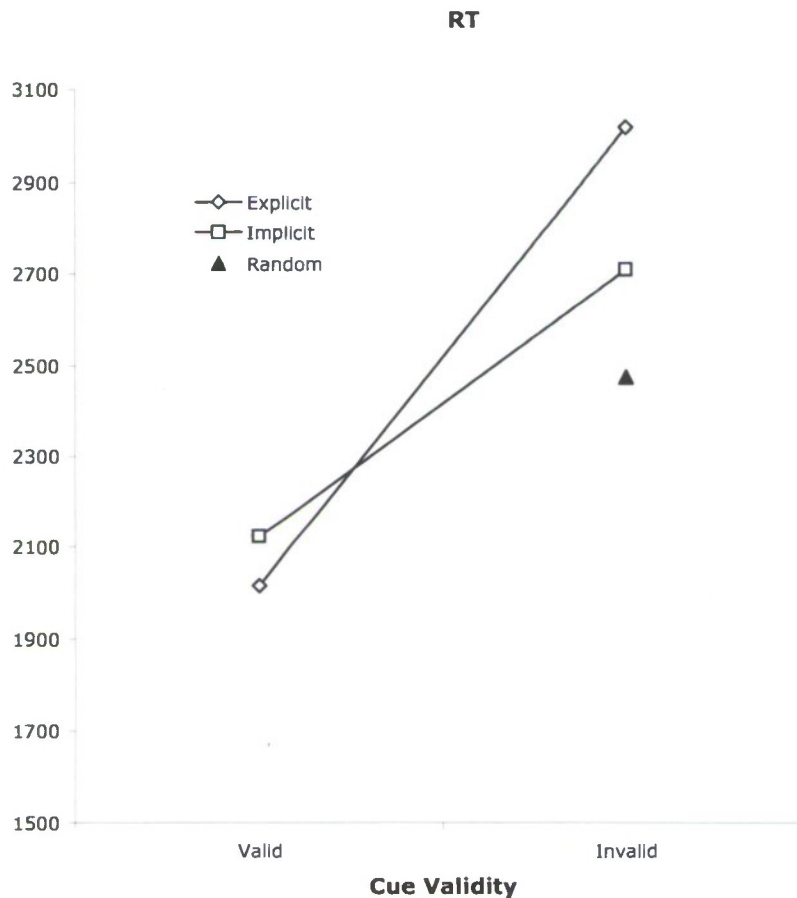
where p_i is the probability that the target will occur at location i , e is the exponential function, and p'_i is the SoftMax prediction for the sampling proportion.

Because of substantial variability between subjects and the small number of subjects the data show only non-significant trends for all comparisons. With that in mind, the patterns do suggest important differences between explicit cueing and implicit learning of likely target locations. In both cases, RT to detect the target increased when the cue (or the likely location) was invalid. This increase was more pronounced for the explicit cue, owing largely to very high detection times in invalid conditions. Sampling behavior also showed intriguing differences. For explicit cues, the total proportion of time spent fixating on the cued location (open diamond symbols in the second figure) did not differ as a function of cue validity. This is what would be expected if that location were simply sampled for frequently, and the observed data is well fit by the simple SoftMax prediction (filled triangles).

For the implicit cueing, where there was one likely location throughout (not designated at the beginning) the pattern shows less than predicted sampling for valid trials, and more than predicted for the invalid case. In other words, subjects preferred exploration on valid trials, exploitation on invalid trials. This pattern is not well fit by the SoftMax function and points to differences in strategy associated with explicit and implicit likelihood. The consideration in interpreting these results is that subjects could not tell a valid trial from an invalid one prior to detecting the target, thus, there could be overt strategy on each from the outset. Differences in fixation duration could have arisen only as a function of what they did during the trial prior to detecting the target. One way to account for the effect of validity on fixation duration in the Implicit condition is to assume that with time sampling became more focused on the more likely position. It may appear to subjects that early on in the trial that they should sample as many locations as they can. This may account for why RT on invalid trials is not as high in the implicit condition, as invalid locations are more likely to be sampled early on in the trial. On invalid trials where the target is not detected early, subjects may switch strategies to sample the likely location as their sampling of less likely locations has not yielded success.

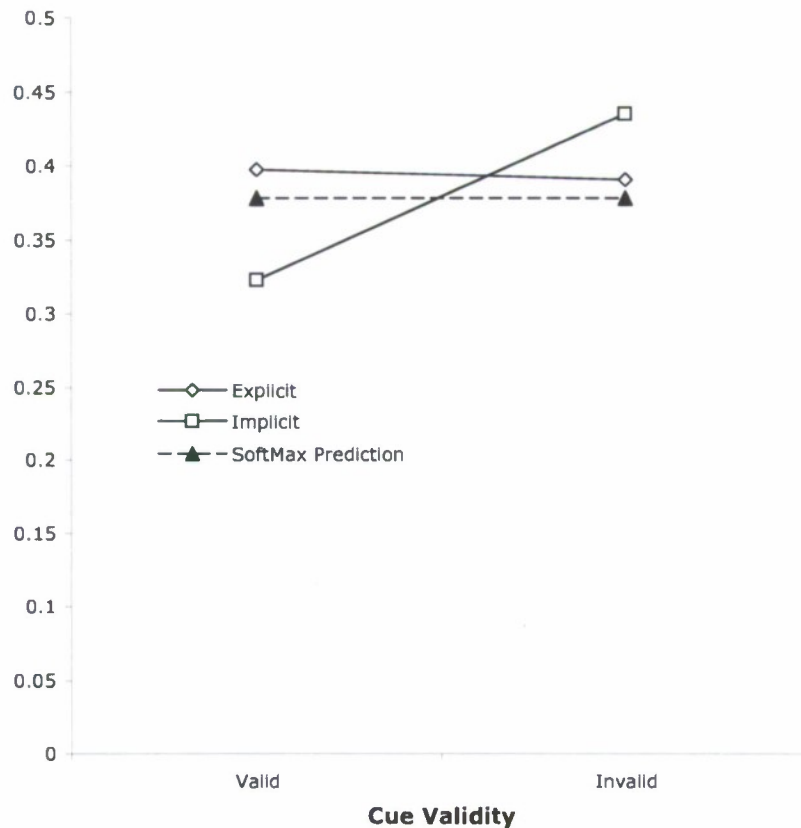
Another possibility is that subjects criterion for target detection changed with time. One of the goals of this paradigm was to have a sufficiently complex discrimination task that subjects would have to integrate information over time to determine which location contained a target. Thus, the location probability could drive both the frequency with which a location was sampled as well as the decision criteria for

detecting a target at a given location. One way to interpret that data then is that the pattern of fixation didn't change but the criteria at the likely location became stricter. On an invalid trial, with time, there would have a repeated number of samples of the likely location without detecting a target. Given that it was sampled more frequently the extra decision time would increase the total fixation duration on the likely location on invalid trials.



At present issues of strategy remain unanswered. Additional subjects and more detailed analyses will be needed to confidently frame an account of behavior. Indeed, one of the advantages of this paradigm is that it was designed to include a significant decision component as well as a location probability component. We have reported these two aggregated as total fixation duration. However, as discussed above they need to be analysed separately to fully understand the strategies employed. At present we have a results suggesting the explicit and implicit probability give rise to different patterns of sampling.

Proportion of Fixations on Cued Location



Conclusions

The research has generated significant findings regard how people sequence their eyes and hand in coordination with task processing to accomplish a series of tasks. These results are now being prepared for submission and are expected to generate important research papers over the next year. Initial investigations of sampling behavior have yielded clues to what could be important differences in the way people use explicit declarative statements of likelihood compared to what they learn through experience. Further experimentation and analysis will be needed to draw firm conclusions from this work.

Personnel Supported

Roger Remington, PhD
Professor
School of Psychology
University of Queensland
St Lucia QLD 4072 Australia

Shu-Chieh Wu, PhD
Senior Research Associate

San Jose State University and NASA Ames Research Center

Harold Pashler, PhD
Professor
Psychology Department
University of California, San Diego

Stefanie Becker, PhD
Post-Doctoral Research Fellow
School of Psychology
University of Queensland, QLD, Australia

Jason Jones
Graduate Student
Department of Psychology
University of California, San Diego

Jessica Irons
Graduate Student
School of Psychology
University of Queensland, QLD, Australia

Publications

Remington, R.W., Wu, S-C, Becker, S., Irons, J. & Pashler, H. (in preparation).
Emergent properties in performing a sequence of discrete tasks: Why the First
Response is so Slow.

Remington, R.W., Wu, S-C, & Pashler, H. (in preparation). What is overlapped in
coordinated eye-hand sequences?

Interactions/Transitions

a. Conference proceedings

Wu, S-C., Remington, R.W., & Pashler, H. (2007). Scheduling of Eye Movements
and Manual Responses in Performing a Sequence of Choice Responses: Empirical
Data and Model. *Proceedings of the 8th International Conference on Cognitive
Modeling*, pp. 331-336, Ann Arbor, MI. July, 2007.

Wu, S-C, Remington, R.W., Lewis, R. (2006) Modeling the scheduling of eye
movements and manual responses in performing a sequence of discrete trials. In,
*Proceedings of the 28th Annual Cognitive Society, Vancouver, BC, Canada, July
26-29.*

Remington, R.W., Lewis, R., & Wu, S-C (2006). Scheduling mental operations in a
multiple-response sequence: Modeling the effects of a strategy to minimize
variance in the timing of saccades. In *Proceedings of the 7th International
Conference on Cognitive Modeling*, Trieste, Italy. April 2-6, 2006.

- b. **Consultative and advisory functions:** Remington invited participant on DARPA panel to evaluate potential programs focused on human machine interfaces to distributed autonomous systems. June 2006, Washington DC.
- c. **Technology assists:** NONE
- d. **Discoveries, patents:** NONE
- e. **Honors:** Remington awarded Australian Professorial Fellowship

Attachment A

Manuscript to be submitted on strategic deferral of RT1 in sequence execution

**Emergent Properties in Performing a Sequence of Choice
Tasks: Why is the first response so slow?**

Roger Remington
University of Queensland

Shu-Chieh Wu
NASA Ames Research Center and
San Jose State University

Stefanie Becker
University of Queensland

Jessica Irons
University of Queensland
&

Harold Pashler
University of California, San Diego

Abstract

In daily life, tasks are commonly accomplished by a sequence of actions, often well practiced, that take a few seconds to complete. These sequences reflect a highly coordinated schedule of overt behavior (e.g. eye movements) and covert cognitive processes. To provide an account of this coordinated activity it is necessary to determine those aspects of performance that arise from fundamental constraints on cognitive processing, as identified in psychological studies, and those that arise from strategic goals not derived directly processing limitations. Across five experiments, eye movements and manual responses were recorded from subjects performing a series of 3-9 choice tasks arrayed linearly on the display. All experiments showed a characteristic pattern of results: elevated response times to the first item, short inter-response intervals with constant dwell times for subsequent items, and a significantly shorter inter-response interval for the final item. The elevated first response was unaffected by sequence complexity, nor was it eliminated by preview, strongly supporting the hypothesis that it is an emergent feature of sequence execution more closely aligned to subject strategies than underlying processing constraints. We demonstrate how the subsequent dwell and inter-response intervals can be derived from a simple underlying bottleneck model of the individual tasks.

Introduction

A central contribution of cognitive psychology has been the demonstration of limits on human information, and a characterization of those limits in terms of a set of processing resources with constraints on how they can be scheduled. For example, when people are required to make independent responses to two tasks presented close together in time, response time to one task is elevated compared to when the task is done in isolation. The generally accepted view is that presenting stimuli for the two tasks closely in time forces overlap in the mental processing of two tasks, resulting in interference when the two tasks compete for limited capacity processing. For a wide range of data from many studies, a simple central bottleneck architecture, similar to that first proposed by Welford (1952; see also, Byrne & Anderson, 2001; Pashler, 1984; Pashler & Johnston, 1989; Ruthruff, Johnston, Van Selst, Whitsell, & Remington, 2003) provides a good account with accurate quantitative predictions. According to central bottleneck theory, three resources -- perception, central processing, and motor execution -- are used to process three successive, independent functional stages, respectively: Stimulus Encoding (SE), Response Selection (RS), and Response Execution (RE). Perceptual processing and motor execution can be done in parallel, but central processing constitutes a single-channel bottleneck. Functionally, this means that SE or RS for a task can be done in parallel with all stages of another task. RS is assumed to be a single-channel bottleneck, so that RS on only one task can be done at any given time (though for an opposing view see Kieras & Meyer, 1997; Meyer & Kieras, 1997a, 1997b).

Processing overlap is not an isolated laboratory phenomenon; concurrent processing of two or more stimuli appears to occur naturally in reading, reaching and grasping, typing, sight-reading music, and other common daily tasks. In such tasks, behavior unfolds as an ordered sequence of overt eye and hand actions. Yet, there is evidence of complex overlapping of the underlying cognitive operations. For example, the eyes are often fixated on a stimulus well ahead of the stimulus being responded to. Such look ahead generally produces inter-response intervals that are faster than response times in isolation. For example, skilled typists fixate several characters ahead of the character being typed, and show short keystroke intervals compared to keypress responses in isolation (John, 1996; Salthouse, 1986); skilled musicians fixate several notes ahead of those being played (Furneaux & Land, 1999), and in reaching and grasping tasks we fixate the to-be-grasped target prior to any movement of the hands (e.g., Epelboim et al., 1993; Epelboim & Suppes, 2001). The inference is that look ahead allows multiple stimuli to be concurrently in different stages of processing: a processing pipeline that promotes rapid and uniform output. Overlap in the processing of successive actions is essential to fluid movement, allowing one action to smoothly blend with those of its neighbours.

How well does the overlap seen in dual-task studies predict the overlap when executing a sequence? That is, how well do our theories of cognitive resources account for performance in executing a sequence of tasks? If constraints on resource allocation were the principle determiner of overlap in sequence execution then it should be possible to predict performance on a sequence from a resource model of each component task, augmented to include the resource demands of shifting from one task to the next. Not only would this be a valuable extension of theory, it would also improve our ability to apply theory to applied problems. Resource constraints

undoubtedly affect workplace performance, workload, and error, but in general, it has proven difficult to generalize from laboratory experiments, with response times on the order 300 – 1200 ms, to performance on daily tasks, which may take 3 - 10 seconds to complete. While there are several reasons for this difficulty, it can be in part that laboratory experiments do not tap the range of strategies people employ, even in executing simple sequences. Laboratory experiments are designed to reduce or eliminate strategies so that architectural features can be seen clearly. Yet, even in a well-defined task, such as reading or typing, people can choose to employ their resources to meet explicit or implicit performance objectives. The central question of the present paper is to what extent the timing of eye and hand events in a sequence execution is determined by fundamental resource constraints or by strategic choices.

Resource Scheduling in Sequence Execution: "Hard Constraints"

Resource constraints have figured prominently in studies of human-computer interaction. Accurate predictions of keystroke and mouse actions in a sequence have been obtained from computational models that combine perceptual, cognitive, and motor demands of each component action with logical dependencies dictated by the task (e.g., the mouse must be over the menu prior to clicking) and by the flow of information (e.g., a stimulus should be perceived before selecting and executing a response). This is true for simple tasks such as withdrawing money from an automated teller (John, Vera, Matessa, Freed, & Remington, 2002; Vera, John, Remington, Matessa, & Freed, 2005), as well as complex applied tasks, such as telephone call handing (Wayne D. Gray, John, & Atwood, 1993; John, 1996). Together, the combination of resources and logical dependencies comprise a set of "hard" constraints (W. D. Gray, Sims, Fu, & Schoelles, 2006) in that they describe fundamental restrictions on processing over which people have no control. The models perform well as engineering approximations, but it can be difficult to verify the specific assumptions that order the perceptual, cognitive, and motor operations.

A study by Pashler (1994) directly tested whether the resource assumptions of a simple central bottleneck could account for manual responses in a sequence. Across several experiments, Pashler had subjects make a 3 (or 4) -choice speeded response to each letter in arrays of 5 – 10 letters presented horizontally on a computer screen. In no-preview conditions the next letter was presented only after the response to the current item was made. In preview conditions, one or more subsequent letters were always present. According to single-channel central-bottleneck theory the RS (central processor) stage limits overlap, since SE and RE can execute in parallel with each other and with RS. The prediction then is that preview allows the RE stage of stimulus N to overlap with the SE (and possibly the RS) stage of N+1, shortening the inter-response interval compared to no preview. This is illustrated in Figure 1.

Insert Figure 1 about here

Across several experiments, Pashler (1994) confirmed this prediction, finding shorter mean IRI with preview than without. Further, varying the luminance of stimuli affected RT1 but had no effect on IRI. This would be expected if SE for N+1 had been done in parallel with RS on N. So long as for a dim stimulus $SE \leq RS$, the effect of luminance on N+1 will be absorbed into the time for RS on N. In contrast, varying stimulus-response compatibility, which should affect the RS stage, showed large effects on both RT1 and IRI, consistent with central bottleneck predictions.

Qualitatively then Pashler (1994) found that the observed IRI satisfied several central bottleneck predictions.

Emergent Properties in Sequence Execution: "Soft" Constraints?

One of the consistent results of Pashler (1994) not so easily derived from central bottleneck considerations, was the large elevation in the response time to the first item (RT1) with preview. Since the first item (S1) cannot benefit from overlap with preceding items the bottleneck model predicts that $RT1 > IRI$. Nonetheless, without additional assumptions central bottleneck theory does not predict a difference in RT1 between preview and no-preview conditions, nor does it predict a difference between RT1 and IRI in the no-preview condition. Yet, both effects obtained: mean RT1 was approximately 225 ms slower in the preview compared to the no-preview condition, while in the no-preview condition, mean RT1 was 150 ms slower than mean IRI. Pashler (1994) attributed the RT1 elevation to "overhead" (or set-up costs) in preparing to perform a coordinated series of events with preview. Overhead preserves resource limits as the principal constraint, in the sense that the extra time for RT1 is the result of additional operations at the beginning of the sequence, some of which would be bottleneck processes. For example, additional processing stages would be needed for *set-up costs* associated with programming the eye movements or manual responses for a sequence. Such set-up costs have been observed in studies of motor learning (e.g., Verwey, 2003), where they have been found to be proportional to the complexity of the ensuing sequence (for a review see, Rosenbaum, 2002). Also, first item responses are often slower even in discrete trial experiments (e.g., Altmann, 2007; Logan & Bundesen, 2003), presumably reflecting preparation, which would lengthen stage durations.

On the other hand, it is also possible that people simply chose to delay the first response, perhaps out of a strategy for coordinating resource allocation in a sequence. A similar elevation of the first response has been observed in typing where it has been assumed to be a strategy designed to get the eyes ahead of the manual responding, setting up a processing pipeline (John, 1996; Salthouse, 1986). The idea that resource allocation is a strategic adaptation to task demands is not new, having been proposed for human monitoring (Dessouky, Moray, & Kijowski, 1995), information search (Pirolli & Card, 1999), block construction (Hayhoe & Ballard, 2005; Gray, et al., 2006), and reading (Reichle & Laurent, 2006). These strategic adaptations to task demands have been referred to as "soft" constraints (W. D. Gray et al., 2006) to emphasize that they are more mutable than the hard constraints. In a sequence, for example, people could decide to try to go as quickly as possible, produce a regular series of manual responses and eye movements, or buffer as many items as they can before beginning to respond, balancing memory load against execution interference.

The choice and use of a strategy represents an active point of control external to the resource theories, such as central bottleneck theory, which lack a well-defined control element. Since strategies are the norm rather than the exception any extension of theory to application domains will be successful only if the strategic adaptation is understood. As the elevation of the first response in a sequence appears to be widespread, we tackle the issue by testing whether the RT1 in a simple sequence of choice RT tasks can be best attributed to resource constraints or to user strategy.

Overview of Present Research

If RT1 elevation results from additional resource demands of set-up costs or other types of sequence initiation overhead then RT1 should be sensitive to factors

that increase or decrease those start-up costs. In experiments similar in design to those of Pashler (1994) we observe the effect on RT1 of factors previously shown to affect start-up costs in motor and eye movement sequences. Eye movements were recorded in addition to the manual responses (RT1, IRI) to better reveal the coordination of underlying processing. Experiment 1 specifically looked at whether the need to make eye movements that are coordinated with hand movements imposes an extra processing cost. Experiment 2 tested the role of preparation in RT1 elevation. Experiments 3 & 4 examined the role of set-up costs by manipulating sequence complexity. Experiment 5 examined whether the elevation is specific to the first item, or to the first response.

To foreshadow, our results show that elevation of first responses cannot be attributed to set-up costs, preparation, or other factors that affect only the first item. Instead, participants appeared to strategically defer the first response until 1 or 2 subsequent items had been fixated. Once the motor output had been initiated a steady-state phase occurred in which the inter-response interval closely approximated the inter-saccade interval established at the outset. This is consistent with a strategy of filling the pipeline before beginning the response output. We simulated sequence execution to demonstrate, in principle, how first item delays could arise from a simple scheduling strategy to deal with resource conflicts between task processing and transitioning from one item to the next. Simulation results indicated that once manual responding began, fixation durations and IRIs followed from resource constraints.

Experiment 1

Experiment 1 examined potential overhead in planning or initiating a series of saccades and coordinated responses. In central bottleneck theory, eye movements should affect RT1 or IRI if they add central processing demands that delay central processing stages of the stimulus. Thus, Experiment 1 tests the demands on central processing imposed by a regular sequence of saccades. Pashler and Carrier [(1993)] found significant dual-task interference with voluntary saccades, but not with saccades generated to a peripheral stimulus. It is unclear whether regular sequences of saccades impose central resource demands at all, on the first saccade, or on each saccade.

In previously reported replications of two experiments in Pashler (1994) we widened the display making a regular sequence of saccades necessary [Wu, Remington & Pashler, 2004; Remington, Wu, & Lewis, 2006; Wu, Remington, & Pashler, 2006]. We found substantially the same mean IRI as Pashler, but a much larger RT1. This additional RT1 elevation could indicate a significant central processing demands for initiating a regular series of saccades. However, Pashler spaced his stimuli approximately 1° apart, far enough that subjects may have made a regular sequence of saccades. To test for demands imposed by regular saccades, Experiment 1 compared a condition where the horizontal extent of the entire sequence of letters was less than 1° to one in which items were centered 5.5° from each other. The size of the letters and their spacing in the wide condition was based on pilot testing showing that it was not possible to accurately identify letters without fixating them. In the narrow condition it is possible to clearly see all letters at once, making it extremely unlikely that participants would make a regular sequence of fixations to each item.

Method

Participants. Sixteen undergraduate students recruited from local colleges and universities near NASA Ames Research Center participated in the experiment, receiving course credit or payment for participation. All participants reported having normal or corrected-to-normal visual acuity.

Apparatus & Stimuli. A Pentium 4 PC controlled the presentation of responses, collection of responses, and storage of data. A separate Pentium 4 computer controlled eye movement recording. Eye movements were monitored with a head-mounted video-based eye tracking system (Applied Sciences Laboratory, Model 501) sampling at 120Hz with a spatial precision of approximately 0.5° visual angle. Eye position was determined by computing the distance between the center of the pupil and corneal reflection of the left eye. Experiments were carried out in a quiet, well-lit room with participants seated approximately 60 cm from a 21" CRT display with a 70 Hz refresh rate used for stimulus presentation.

The primary stimulus display consisted of a row of five letters centered at the middle of the display. Each letter subtended 0.34° in height, presented at a luminance of 11.7 cd/m². In the wide spacing condition, the letters were spaced approximately 5.5° apart. In the narrow spacing condition, the whole span of the letters subtended less than 1° of visual angle. Stimuli were the letters T, D, Z presented in uppercase to which participants responded by pressing the V, B, N keys, respectively, on standard computer keyboard.

Procedure. The experiment consisted of a total of 120 trials divided equally into two blocks, with one block for each spacing condition. The order of the two spacing conditions was counterbalanced across participants. Half of the participants received the narrow spacing condition first, followed by the wide spacing condition. The other half of the participants received the reversed order. Prior to the experiment each participant completed 24 practice trials of the first condition assigned.

Each trial began with the presentation of a white fixation cross (0.3°) in the center of the display. After the participant had maintained fixation within a 6° radius around the fixation for 500 ms, the fixation was erased and a small filled square (0.34°) appeared at the leftmost stimulus position. Participants were instructed to fixate the small square and maintain fixation until the stimuli were presented. The small square remained for 1 sec, followed by a blank interval of 500 ms, after which the 5 stimulus letters were presented. Eye movement recording began the moment the small square appeared over the location of the leftmost item, and ended after the participant had responded to the rightmost stimulus. A calibration procedure was administered before each block of trials to maintain accuracy of recordings. The characters were erased after the participant had responded to the rightmost character. The next trial began following an inter-trial-interval of 250 ms.

Participants were given a written description of the task, which was reviewed with the experimenter. They were instructed to respond to each item as quickly and accurately as they could and not to group their responses.

Manual responses and eye fixations for each item were recorded. Eye fixation samples were analysed offline to classify them into saccades or fixations, and assign fixations to stimuli. Because the stimuli were arrayed horizontally at the same vertical screen position, all analyses were based on horizontal (x-axis) movements only. A saccade was defined as a movement velocity exceeding $30^\circ/\text{s}$ or movement acceleration exceeding $3000^\circ/\text{s}$. A fixation was defined as movement velocity below $30^\circ/\text{s}$ or movement acceleration fell below $-3000^\circ/\text{s}$. A fixation was assigned to the

nearest stimulus letter position and its duration was calculated by summing all contiguous individual fixations on a designated target region. Once a fixation on an item ended subsequent fixations on that item were considered regressions. Fixations above or below the stimulus array, or to either side of it, were considered anomalous and omitted in the analyses.

Results

Sequences containing regressive fixations or fixations outside the letter sequence were excluded; manual and eye fixation results represent only those trials that contained a clear sequence of left-to-right eye movements, interpretable in terms of the task. Analyses include only items correctly responded to. Mean correct responses time on S1 (RT1) and mean correct Inter-Response Interval (IRI) for items S2-S5 were computed for each subject.

Three measures were computed for the eye fixation data. Eye-Hand Span (EHS) is the time from the initial fixation on a stimulus till its response. In isolation, the EHS is equivalent to RT. In a sequence, the overlap in processing of adjacent items means that the EHS may reflect postponement of task processing to include operations on previous and subsequent stimuli. Note that our use of the term "Eye-Hand Span" differs from earlier studies, which used it to refer to the number of items ahead the eyes were when the response was made. Dwell time is the duration of fixation on a stimulus, and our usage corresponds to its common usage. Release-Hand Span (RHS) is the time from when the eyes leave stimulus N (presumably to fixate stimulus N+1) to when the response to stimulus N is made. RHS is derived from the other two eye movement measurements by subtracting Dwell from EHS. RHS relates directly to the overlapping processing of two adjacent stimuli, since processing on

Insert Figure 3 about here

stimulus N is still in progress while the eyes are fixated on N+1.

Figure 2 is a graphic representation of the observable measures from Experiment 1 annotated to illustrate the manual and eye movement measures. Rows represent successive stimuli S1-S5 from top to bottom, with time running horizontally. The bar for each row represents the mean EHS for its corresponding stimulus. The shaded portion is the mean Dwell time; the unshaded portion is the

Insert Figure 2 about here

RHS. RT1 corresponds to the EHS for S1, while IRI measures the time interval between successive responses. As Figure 2 shows there is considerable overlap between eye fixations and manual responses.

For manual responses, analyses were conducted separately on mean response times to S1 (RT1) and the mean Inter-Response Interval (IRI) for S2-S5. The left hand panel of Figure 3 plots mean RT1 and IRI for Narrow and Wide conditions. Data from Figure 5 of Pashler (1994) is included for comparison. Mean RT1 in the

Narrow condition was 1264 ms compared with 1182 ms for the Wide. This difference was not significant by a paired 2-sample t-test ($t = -1.19$, $df = 15$, $p = .252$). A repeated measures analysis of variance was conducted on the mean IRI for each subject with width (Wide, Narrow) and stimulus (S2-S5) as factors. The main effect of width was significant ($F[1, 15] = 8.24$, $p < .015$), as was the main effect of stimulus ($F[1, 15] = 11.756$, $p < .001$), and the interaction of width and stimulus ($F[3, 45] = 3.687$, $p < .02$). Mean IRI in the Wide condition was 417 ms compared with 376 ms for the Narrow. Post-hoc t-tests showed the interaction of stimulus and width due to significant shorter IRIs for the Narrow condition at S4 and S5.

The eye movement data is shown in the right hand panel of Figure 3. It is clear that both the EHS and RHS decline sharply over the first three stimuli while Dwell remains constant. An analysis of variance on Dwell times as a function of stimulus revealed no main effect of stimulus ($F[1, 15] = 1.94$, $p > .11$). The slope of Dwell against stimulus number was -7.7 ms per item. Corroborating the flat slope of Dwell, the correlation of EHS and RHS was .99.

Discussion

The qualitative patterns of RT1 and IRI for wide and narrow displays were very similar, suggesting that the need to make eye movements had little if any effect on resource scheduling. Narrow displays had a marginally larger mean RT1, while wide displays resulted in slightly longer mean IRI. The results provide no support for set-up costs in sequence initiation as the source of RT1 elevation. Experiment 1 also shed light on another observed feature of sequence data. Here, as well as Pashler (1994) and our earlier replications, mean IRI was between 400-500 ms, surprisingly high for an estimate of the RS stage in such a simple task. This suggests that transitioning from one stimulus to the next does impose a central demand that is more associated with shifting attention, or task set, than with the saccade per se. Note too that the IRI for the final item was about 100 ms shorter than the previous three. The key difference is that the final item requires no further transition. This "last-item" effect in IRI does not appear on the graph of Pashler's data. That data was taken from the first 5 stimuli from a longer series, so that S5 was not the final stimulus. The last-item effect suggests that there is a resource conflict between transitioning to the next item and processing of the current one, resulting in a 100 ms delay either in RS or RE for the task. A central RS duration of about 300 ms would be much more in line with dual-task studies (Pashler, 1984).

Experiment 2

Experiment 2 tested a different possible source for RT1 elevation, the extent to which it results from inefficient processing of S1 itself. First trial costs have been observed in studies of task switching (Altmann, 2007) and it is not uncommon to discard the first trial of a block in discrete-trial experiments as participants may not be fully prepared prior to completing response selection for at least one stimulus. Studies of task switching have found a residual switch cost with even long response-stimulus intervals suggesting that preparation is not complete until task processing (presumably response selection) has been completed (see e.g. Rogers & Monsell, 1995; Ruthruff, Remington, & Johnston, 2001). In Experiment 2 we attempted to eliminate any possible inefficiency in S1 processing by including a condition where participants were allowed a preview of S1 long enough to fully process it and retrieve

any sequence-related plans. When the sequence began then they should have been fully prepared to respond to S1.

Method

Participants

Six participants, four males and two females, from The University of Queensland, Australia, took part in the experiment as paid volunteers (\$10/h). Mean age of participants was 34.17. All subjects had normal or corrected-to-normal vision and were naive as to the purpose of the experiment.

Apparatus, Materials & Stimuli

Computers, eye tracking equipment, and displays were identical to those in Experiment 3. The fixation display consisted of a row of 7 figure-eight filler characters; target letters (H, S, U) were revealed by offsetting elements of the filler characters. Fillers and letters measured $0.20^\circ \times 0.25^\circ$ and were evenly spread over a wide viewing area (26.8°) in the centre of the display. The distance between the centers of two adjacent characters measured 4.67° , with the outer stimuli being 3.2° apart from the monitor frame. The presentation of stimuli, collection of responses, and timing of events were controlled by an Intel(R) 2CPU 2.4GHz-Computer (Dell) with a 21" SVGA colour monitor (BenQ). The experiment was controlled by the Presentation software package (Neurobehavioral Systems). Stimuli were presented with a resolution of 1,280 x 1024 pixels and a refresh rate of 99.9Hz. Eye movements were recorded using an EyeLink1000 with spatial resolution of 0.05 and a tracking rate of 250Hz. Participants were seated in a dimly lit room, with their head in a chin rest with forehead support, at a distance of 64 cm. from the screen. All stimuli in the experiment were presented in black (RGB: 0, 0, 0) against a dark grey background (RGB: 100, 100, 100). In the no-preview condition, the initial fixation display consisted of 7 fillers; in the preview condition the leftmost target, S1, was exposed with the remaining 6 being fillers. Figure 4 shows an example of the stimuli in each condition of Experiment 1.

Insert Figure 4 about here

Design & Procedure.

Preview and No-Preview conditions were presented in separate blocks counterbalanced for order of presentation. Participants completed 30 practice trials, which were not recorded, followed by 180 experimental trials in each block. Each trial started with the presentation of the fixation display, which consisted of 7 filler items in the no preview condition, and the S1 letter with 6 filler items in the preview condition. In all other respects, procedure was identical to that of Experiment 3.

Results.

In Experiment 2 we excluded all trials in which one or more items had been responded to incorrectly. This amounted to a loss of 11.11% of data in the no-preview

condition, and 9.35% in the preview condition. Trials were also excluded when the mean IRI was above 4,000 ms, which affected fewer than 0.01% of trials. Data were subjected to a 2 x 7 repeated measures ANOVA with variables of Preview (with preview vs. no preview) and position in sequence (1 to 7). Two separate 1 x 7 ANOVAs tested for sequence effects. For all analyses, the Greenhouse-Geisser corrected p-values are reported, together with the uncorrected degrees of freedom.

As shown in Figure 5, mean RT1 was 553 ms slower than the mean IRI of the remaining letters. This difference was significant both for no-preview ($F(6,30) =$

Insert Figure 5 about here

37.99; $p < .001$) and preview conditions ($F(6,30) = 20.23$; $p = .001$). There was a significant interaction between sequential effects and preview condition ($F(6,30) = 10.81$; $p = .010$). A 2-tailed paired t-test ($t = 2.93$, $df = 5$, $p < .05$) showed that the source of this interaction was a slower mean RT1 for no-preview (953 ms) than for preview to (800 ms). Mean IRI averaged over position did not differ significantly between the two preview conditions (2-tailed paired t-test, $p > .27$).

A 2-tailed paired t-test showed a marginal effect of preview on Dwell time ($t = 1.98$, $df = 5$, $p < .15$); mean Dwell in the No Preview condition was 498 compared with 417 in the Preview condition. As shown in Figure 5, Dwell times for S2-S7 did not differ significantly between preview and no-preview conditions (all $ps > .26$). The effect of preview on Dwell was 80 ms compared to 160 ms for RT1.

Analysis of variance on mean EHS showed no significant main effect of preview condition ($F < 1$), a main effect of letter position ($F(6,30) = 15.04$; $p = .007$), and a significant interaction between preview condition and letter position ($F(6,30) = 13.59$; $p = .002$). The interaction resulted from longer RT1 in no-preview condition than in the preview condition reported above, with no effects of preview on mean EHS for subsequent letter positions (all $p > .14$).

Discussion.

RT1 decreased with 1.5 seconds of preview, from 950 ms (no-preview) to 800 ms (preview), evidence that subjects used the preview to shorten S1 processing time. Nonetheless, 800 ms is a far higher response time than would have been expected for a single stimulus-response trial without preview. Pashler (1994) reported a 3-choice RT1 of 700 ms when each item was done in isolation (i.e., S2 was not revealed until after the response to S1). Here in Experiment 2, where the next item could be previewed, participants took about 100 ms longer than in Pashler (1994) to respond to a letter they had been looking at for 1.5 seconds. In addition, R1 occurred only after having fixated S2 for 400 ms, only slightly less than without preview. In short, the reductions in RT1 with preview are much less than would be expected had preparation or S1 processing delays been the source of RT1 elevation. The cause of the elevated first response is not to be found in S1 processing per se. Nor is it a simple function of doing items in sequence; Pashler (1994) observed only small increases in RT1 in conditions where subsequent items were not present. Rather, the elevation of the first response emerges from the attempt to overlap the processing of adjacent items in sequence.

Preview had two dissociable effects on RT1, which could provide clues as to how sequences are organized. First, preview reduced Dwell from about 500 ms to 415 ms with preview. Secondly, it reduced RHS from 470 ms to 400 ms. Dwell has been found to be affected by the difficulty of stimulus processing (see e.g., Reichle, Pollatsek, Fisher, & Rayner, 1998; S.-C. Wu, & Remington, R. W., 2004; S.-C. Wu, Remington, R. W., & Pashler, H., 2004). The reduction in Dwell then is consistent with S1 processing during the preview interval. Less is known about what processes occur in the RHS epoch, the portion of processing remaining after the eyes have moved. The effect of preview on RHS could indicate that RS (Response Selection) is not complete when the eyes move. In bottleneck theory only RE (Response Execution) remains after RS is complete. It is unlikely that the 70 ms effect of preview was due to speeding of a simple key press. Instead, some or all of both effects should be attributed to a reduction in RS. If so, this would suggest that saccade initiation is not dependent on the completion of RS.

What is clear from Experiment 2 is that RT1 is determined in large measure by how the response to S1 is coordinated with other activities in the sequence. What remains a puzzle is why after previewing S1 it took 400 ms before making the saccade to S2, and why a further 400 ms was needed to respond to S1. This could signal costs in programming the coordinated eye-hand sequence, or delays due to resource conflicts in scheduling the eyes and hands with internal processing. In the remaining experiments we address what role the nature of the sequence plays and whether there is a deliberate strategy of deferring R1 (buffering) or whether the delay is a direct reflection of resource conflicts.

Experiment 3

One way to affect the programming time for a sequence is to increase the complexity of the programming. For learned motor sequences, time for the first response is roughly proportional to the complexity of the sequences, as measured by sequence length, heterogeneity of items, and heterogeneity of item spacing (see Rosenbaum, 2002). Yet, a comparison of Experiments 1 and 2 shows that increasing the number of items from 5 to 7 had little effect on the pattern of results. Likewise, Pashler (1994) tested list lengths of 5, 10, and 6 items across experiments with no systematic effects of length on RT1. Our sequences, and those of Pashler, were very regular with length held constant for an entire experiment. In Experiment 3 we varied sequence length from trial to trial, making it impossible for participants could not prepare for a predictable sequence length. In Experiment 3a, 9 locations were always presented and sequence length was varied by inserting filled rectangles as placeholder objects after 3, 5, or 9 items, keeping inter-item spacing constant. In Experiment 3b, lists of 3, 5, and 9 items were presented keeping the total extent identical, with item spacing varied unpredictably across trials.

Method

Participants & Apparatus

For Experiment 3a, 12 undergraduate students from the same NASA Ames Research Center subject pool as Experiment 1 participated in the experiment. All met the same requirements specified in Experiment 1. The apparatus used in Experiment 2a was identical to that in Experiment 1. For Experiment 3b, 18 undergraduate

students recruited from introductory psychology classes at the University of Queensland participated in the experiment.

Stimuli and Display

For Experiment 3a, the primary stimulus display consisted of a row of 9 letter positions, spaced approximately **xx** apart, centered on the middle of the display. The stimulus letters on each trial were aligned with the leftmost position, with the rest of the positions occupied by small filled squares. The letters T, D, & Z were assigned to the V, B, and N keys, respectively, of a standard keyboard.

For Experiment 3b, characters were evenly spaced across the 9 possible character positions with the first and ninth positions always occupied. Thus, for a sequence of 3 items, positions 1, 5, and 9 were occupied. For sequences of 5, positions 1, 3, 5, 7, and 9 were occupied. The letters T, D, & Z were assigned to the V, B, and N keys, respectively, of a standard keyboard.

Design and Procedure

The design and procedure were identical for Experiments 3a and 3b. Each experiment consisted of 180 trials, 60 in each sequence length condition, administered in 3 blocks of 60. List length varied from trial to trial within a block. Prior to the experiment participants received 24 practice trials consisted of all three types of sequence length. Trial sequence followed that described in the previous experiments.

Results

Experiment 3a

As in Experiment 1 we analysed only data for correct items from sequences with regular saccades. The open symbols in the leftmost panel of Figure 6 show manual responses in Experiment 3a for sequences of 3, 5, and 9 items. As is clear from Figure 6 there was no difference in mean RT1 or IRI between sequences of different length. This was confirmed by t-tests on pairs of means at S1 and S2 with $0 < t < .2$ in all cases.

The middle panel of Figure 6 shows Dwell and EHS for Experiment 3a. Paired t-tests found no significant effects of sequence length, $0 < t < .2$, for any comparison. Error rates were uniformly low (below 2% for all stimuli) and were not subjected to further analysis.

Experiment 3b

Filled symbols in the leftmost panel of Figure 6 show manual responses in Experiment 3b. Mean RT1 for 3, 5, and 9 items was 1076, 1114, and 1155 ms, respectively. Paired t-tests on corresponding means showed none of the contrasts

Insert Figure 6 about here

approached significance ($0 < t < .5$ in all cases). Overall mean RT1 collapsed across the three list lengths was faster in Experiment 2b than in Experiment 2a (paired t-test for unequal samples, $t = 18.84$, $df = 28$, $p < .01$).

The rightmost panel of Figure 6 shows mean EHS and Dwell time for Experiment 3b. Dwell for each sequence length were virtually identical with each other with those from Experiment 3a. EHS shows a trend toward increasing EHS with increasing number of stimuli. Mean EHS for sequences of 3, 5, and 9 items were 1084, 1113, and 1161 ms, respectively. Paired t-test on corresponding means found no significant effects ($0 < t < .2$, $df = 34$ for all cases). As in Experiment 3a, error rates were uniformly low (below 2% for all stimuli) and were not subjected to further analysis.

Discussion

Increasing the number of items was not associated with either increases in mean RT1, IRI, EHS, or Dwell. The results of Experiment 3b show that the absolute distance between items had no effect on performance even though the spacing was not known until the display came on. Unlike motor sequence execution, where set-up costs are a function of number of items, RT1 here showed no effect of number of items. A likely explanation of why length had no effect in our sequences is that an independent judgment was required on each stimulus in turn. Unlike learned motor sequences, or eye movement patterns, there is no inducement to treat the entire sequence as a unit. Another consequence of the independent choice tasks is that the Dwell times and IRI cannot be shorter than the time required to make the choice decision. If the goal were only to make rapid eye movements or manual responses set-up costs to facilitate the sequence might have been observed. Nonetheless, the conclusion is clear: RT1 elevation cannot be attributed to set-up costs similar to those observed with learned or cued motor sequences.

The 9-item display did reveal patterns not present in sequences of 5 or 7 items. With 9 items it is clear that EHS declined over the first two items, remaining relatively flat from the third item onward. The first phase in executing a sequence begins with a regular sequence of eye movements prior to any motor output. In this first phase, then, eye movements are decoupled from manual responses. This is followed by a second, steady state phase with a constant timing relationship between the saccade initiation and manual response. In fact, from S2-S8 mean Dwell was very close to mean IRI, 532 ms compared to 520, respectively. Deviations from the equivalence of IRI and Dwell occur only for S1 and S9, where first and last item effects, respectively, alter the pattern. How this coordination arises is not yet apparent. It is consistent with a strategy of postponing R1 to allow the "pipeline" to fill prior to beginning the manual response sequence, with saccade rate adjusted to match the expected manual output rate.

We speculated earlier that the reduced IRI immediately preceding the last item, the last-item effect, occurred because there was no need to coordinate the transition to the next item. As a result, RS stages of task processing did not conflict with central processor stages for the transition. Further support for this comes from the 3- and 5-item displays in Experiment 1a. In Experiment 1a all 9 possible locations were occupied, with placeholders filling out the remainder for sequences of length 3 and 5. Thus, subjects did not know they were at the end of the sequence until they had fixated on the first placeholder. Thus, the transition costs are incurred on the last item. In Experiment 1b, where it is clear that there is no subsequent item, there was a last-item reduction of IRI of similar magnitude to that found in earlier studies.

Experiment 4

Thus far, sequence complexity has had little effect on RT1. Still, the sequences were regular in spacing for any given trial. If set-up costs for sequence initiation were done at the beginning of each trial then Experiment 3b tested only whether set-up costs were greater for larger spacings. Experiment 4 compared regularly and irregularly spaced stimuli within a trial to measure any differences in set-up costs as a function of display complexity. The logic was that regular, predictable spacing should be less complex and require less set-up time than conditions in which spacing varied unpredictably within a trial.

Method

Participants & Apparatus

Six participants (three male, three female) recruited from The University of Queensland, Australia, took part in the experiment as paid volunteers (\$10/h). Mean age of the participants was 34.67 years. All subjects had normal or corrected-to-normal vision and were naive as to the purpose of the experiment. The Apparatus was identical to that in Experiment 3b.

Stimuli, Design and Procedure.

Stimuli consisted of the letters H, S and U mapped to the V, B, and N keys of a standard keyboard. All stimuli in the experiment were presented in black (RGB: 0, 0, 0) against a dark grey background (RGB: 100, 100, 100). The fixation display consisted of a black box of the same size as the letters ($0.20^\circ \times 0.25^\circ$) centered on the leftmost position of the display. Target letters were presented when the tracking was stable and the gaze was fixated on the centre of the black box (within 1.3°), for at least 500 ms, within a time-window of 3,000 ms. Each stimulus display contained five letters distributed across 9 equally spaced possible character locations. The letters were drawn randomly from the stimulus set H, S and U with the restriction that none of the letters was drawn more often than twice and that the same letter could not appear twice in a row. The distance between two adjacent potential target locations measured 3.7° , and the middle position was centred on the display.

In the regular spacing condition, letters were presented at positions 1, 3, 5, 7 and 9, so that the distance between two adjacent stimuli always measured 7.4° . In the irregular spacing block, the five letters were randomly assigned to the 9 possible positions, with the restrictions that the first and last positions were always occupied by a letter, and the restriction that spaces of different magnitude (i.e., 2 spaces (7.4°), 3 spaces (11.1°) and 4 spaces (14.8°)) were drawn with equal probability. A trial was terminated when the program detected 5 key presses. After each trial, participants received feedback in the same way as in earlier experiments.

The regular and irregular conditions were blocked, with the order of blocks counterbalanced across participants. Participants received 30 practice trials drawn randomly from the first block before testing commenced. Each participant completed 300 trials, 150 trials in each condition and was allowed a short rest between the two blocks. On average, it took 30 minutes to complete the experiment.

Procedure

Procedure followed that of previous experiments with minor changes. The presentation of the stimulus display was contingent upon the gaze position and was presented only when the tracking was stable (no blinks) and the gaze was within 50 pixels (1.3°) of the centre of the leftmost stimulus. When participants had fixated for

at least 1,500 ms on the leftmost placeholder stimulus (within a time window of 3,000 ms), a short tone sounded (50 ms, 700 KHz). Simultaneous with the onset of the tone, the 5-letter stimulus display was presented. If fixation was not maintained, participants were calibrated anew and the trial started again with the fixation display. Before each block, participants were calibrated with a 9-point calibration and were given written instructions about the next block.

Results

As in Experiment 2, all trials in which an error was made to one or more items, as well as trials with irregular patterns of eye movements as described above, were excluded from analysis. This amounted to a loss of 10.55% of the trials in the irregular spacing condition, 9.44% in the regular spacing condition. Additionally, trials in which the mean total completion time exceed 4 seconds were excluded, resulting in a loss of 0.13% of all trials.

Irregular versus Regular Spacing

Mean RT1 and IRI, shown in the left panel of Figure 7, did not differ for regular and irregular spacing. A 2 x 5 ANOVA with factors of spacing (regular vs. irregular) and position of letter in sequence (1 to 5) found no main effect of spacing ($F < 1$), a significant main effect of letter position ($F(4,20) = 46.81$; $p < .001$), and no interaction of spacing and position ($F(4,20) = 2.32$; $p = .14$). The main effect of position reflects the elevated RT1 compared to IRI for subsequent letters. On average, mean RT1 was 552 ms slower than the mean IRI of the subsequent letters in the regular spacing condition, and 530 ms slower in the irregular spacing condition. Mean RT1 for the regular spacing condition was 962 ms compared to 938 ms for the irregular. This difference was not significant ($t = .765$, $df = 5$, $p < .48$).

Mean Dwell times, also shown in the left panel of Figure 7, did not differ between the regular and irregular spacing conditions ($F < 1$). Dwell times were slightly elevated on the first letter, by 84 ms in the regular spacing condition and by 93 ms in the irregular spacing condition, but the main effect of letter position failed to reach significance ($F(4,20) = 2.73$; $p = .12$). The two variables did not interact with each other ($F < 1$).

Mean Eye-Hand Span (EHS) and Release-Hand Span (RHS) are shown in the right panel of Figure 7. An analysis of variance on mean EHS in each of the conditions showed no significant effect of spacing ($F < 1$), a significant effect of the item ($F(4,20) = 42.19$; $p < .001$), with no spacing by item interaction ($F = 1.35$; $p = .30$). EHS showed a decline of around 250 ms over the first three items, a 21 ms decline between S3 and S4, followed by a 124 ms decline on S5. This is a similar pattern to that of Experiment 3 supporting the idea of an initial phase of sequence execution that transitions into a steady state phase. The decline on S5 again shows the last-item effect characteristic of sequence processing in this paradigm. The reduction in EHS for the last item is consistent with the hypothesis that EHS for other items includes processing of surrounding items and the transition between items.

Analyses of Individual Spacings

To examine the effect of inter-item spacing in more detail performance was analysed as a function of the distance from preceding or subsequent characters. For these analyses, RT1 was excluded and the remaining 4 items analysed as a function of number of distance (number of blank spaces) immediately preceding and following an

item. Mean IRI increased linearly with distance from preceding item ($F(4,20) = 15.05$; $p = .002$), but this was not true for any of the eye movement measures; Dwell, EHS and RHS were unaffected by the number of spaces preceding the responded-to-letter ($F < 1$; $F(4,20) = 1.18$; $p = .35$ and $F < 1$, respectively). The IRI effect appears to be a consequence of more frequent corrective saccades for larger spacings. To test this, we classified saccades as corrective when their starting points and endpoints were located within the region of 1.3° from the centre of the letter. Corrective saccades were classified as undershoots when the endpoint of the initial saccade into the region was located to the left of the target letter, and as overshoots being when to the right. Corrective saccades were significantly more frequent when the responded-to letter was preceded by larger spaces ($F(4,20) = 8.2$; $p = .009$). Thus, the timing of the eye is the same for small and large spacings, but the accuracy of the eye movement is reduced for larger spacings resulting in more corrective saccades that increase IRI. No significant effects were observed in any of the measures as a function of the amount of distance of the letter subsequent to the target.

Discussion

Regularity of spacing had no effect on any of the measures of interest. This differs from reading where spacing irregularity has been shown to increase mean dwell time (e.g., Reichle et al., 1998). Once again, our lists of choice tasks failed to show eye and hand patterns found in other sequence tasks. This contextualization would be expected if eye movement dynamics were adapted closely to task demands. Nonetheless, if RT1 reflects set up costs for the sequence they are not a function either of the number of items, as shown in Experiment 3, nor the regularity of spacing.

It is possible that complexity lies not with the layout, but instead with the demands of the resource scheduling itself. Thus far, our sequences required the eye movements to be sequenced with manual responses beginning with the first stimulus. Would there be a similar elevation of RT1 if the first response occurred not on S1 but on S2? Arguably, by S2 set-up costs for sequence initiation would already have had their effect. Elevation of RT1 when it occurs on S2 could reflect a cost for coordinating the manual response into the eye movement sequence, or strategic buffering of responses.

Experiment 5

To isolate the demands of eye-coordination from set-up costs in sequence initiation Experiment 5 included 2-3 no-go stimuli in each sequence. For half the sequences the first trial was a Go stimulus, for the other half a No-go stimulus. When the first stimulus was a No-Go, the second was always a Go. Set-up costs should effect RT1 only when the first stimulus was responded to. If set-up costs for sequence initiation were the source of RT1 elevation, then no elevation should be seen when RT1 occurred on the second item.

Experiment 5 also included two control conditions. In the Respond-Only (RO) condition participants were instructed simply to respond to the first item in the sequence, ignoring the others. In the Respond-then-Scan (RTS) condition they were instructed to fixate all items but respond only to the first item. RO provided a manual response baseline free from any overhead of making saccades. RTS provided a manual response baseline that included the overhead of make saccades, but without the requirement to coordinate the two on each item. RT1 would still reflect initial

sequence preparation costs for item transition (i.e., saccades, attention shifts), but no need to buffer responses since no further responses would be required. If the elevated RT1 reflects the overhead in coordinating manual and oculomotor responses, then RT1 should be fast in this condition.

Method

With the exceptions noted below the method of Experiment 5 followed that of the wide spacing condition of Experiment 1. Fourteen undergraduate students recruited from local colleges near NASA Ames Research Center participated in the experiment for course credit. The experiment was conducted using a PC with a 21-inch monitor. Participants were seated in a comfortable chair with their head secured on a head-and-chin rest placed 53.5 cm in front of a 21-inch CRT monitor. Eye movements were recorded with an infra-red video-based eye tracking system (ISCAN) with an output rate of 120 Hz.

There were three conditions: Sequence, Respond-then-Scan (RTS), and Respond-Only (RO). The Sequence condition consisted of a 5-item sequence that included either 2 or 3 no-go stimuli. Six lists were constructed that differed in the number of responses (one, two, or three), and in the stimulus position on which the first response occurred (first and second). The six lists can be represented as: TXXTT, TTXXT, TTTXX, XTXXT, XTTXX, and XTITX, where T denotes a target (go) stimulus that required a key press, X a non-target (no-go) stimulus. Go stimuli were randomly drawn from the letter set T, D, and Z, with the constraint that no letter was repeated in two adjacent positions. This constraint however does not prevent repetition of responses; the same letter could occur in two positions separated by an interposed X. The no-go stimulus was the hash character, "#", matched in vertical and horizontal extent with the three go characters. Five participants completed 40 trials of each type administered in 2 blocks of 120 trials. Nine participants completed 60 trials of each type administered in 3 blocks of 120 trials.

The two control conditions, Respond-Scan (RTS) and Respond-Only (RO), consisted of a single target (Go) stimulus in the first position (i.e., TXXXX), differing only in instructions. In the RS condition, participants were instructed to respond to the first letter stimulus then fixate each subsequent item in turn. In the RO condition (i.e., T___), they were instructed only to respond to the first stimulus. There were 40 trials in each control condition. The two control conditions were administered after the experimental conditions and in the same order (Respond-Scan first, Respond-Only second) to each participant.

As before, participants were instructed to respond quickly but accurately. No single aspect of task performance (e.g., manual or oculomotor, speed or accuracy, etc) was emphasized. The only specific instruction given to the participants was to treat each character independently and not group responses.

Results

As in Experiment 1, items responded to incorrectly were excluded from analysis. The left panel of Figure 8 shows manual responses for each sequence separately. The introduction of no-go stimuli meant that not all target processing was equivalent. For example, the response to S5 in TTXXT follows two no-go stimuli as opposed to TXXTT. Consequently, mean IRI was computed only for pairs of immediately successive go (T) stimuli. The left panel of Figure 8 shows mean IRI for

these pairs was fast across all stimulus positions when compared to RT1, or to a target preceded by a no-go stimulus. Analysis of variance on IRI showed that the decline in IRI from S2-S5 was significant ($F[3,39] = 4.57, p < .01$).

Mean RT1 when it occurred on S1 was 890 ms compared to 794 ms when on S2, which was significant on a 2-tailed paired t-test ($t = 5.14, df = 13, p < .001$). To determine whether RT1 on S2 was itself elevated we compared it to the EHS for the second response (S4) on sequence TXXTT. Since S4 in this sequence occurred after two successive no-go stimuli its EHS would be free of effects of previous stimuli. A t-test found RT1 on S2 slower than EHS on S4 ($t = 6.26, df = 13, p < .01$). The final responses in XTXXT and TTXXT were not used because of confounds with last-item effects. RT for the two control conditions, RTS (Respond-then-Scan) and RO (Respond-Only) were 646 and 576 ms, respectively, faster than RT1, and significantly different from each other ($t = 2.97, df = 13, p < .02$).

The right panel of Figure 8 plots shows mean EHS (Eye-Hand Span) for each item in each sequence along with mean Dwell time for go (T) and no-go (X) stimuli averaged over sequences. Analysis of variance on mean Dwell found no effect of position ($F < 1$). Dwell was affected by whether it was preceded by a Go or No-Go stimulus. An analysis averaged over position with preceding and current stimulus type (go, no-go) as factors found significant main effects of current stimulus type ($F[1, 13] = 21.15, p < .001$), previous stimulus type ($F[1, 13] = 41.60, p < .001$), and an over additive trend in the interaction ($F[1, 13] = 3.63, p < .10$). Mean Dwell on targets (go) preceded by a non-target (no-go) was 400 ms compared to 460 ms when preceded by a target. Mean Dwell on non-targets preceded by a non-target was 312 ms compared to 400 ms when preceded by a target. When S1 was the first target Dwell on S1 was 460 ms, compared to 400 ms Dwell on S2 when S2 was the first target ($t = 3.75, df = 13, p < .01$).

For Eye-Hand Span, Figure 8 shows a near linear decrease in EHS over the first three stimuli in sequence TTTXX, and is closely matched over the first two stimuli by TTXXT. EHS (RT1) for the first response in sequences TXXTT, TTXXT, and TTTXX, was unaffected by subsequent stimuli; paired t-tests showed no significant difference between sequences ($t = 1.74, df = 13, p > .10$ for the extreme comparison). A paired t-test on R1 for sequences XTXXT, XTXX, XTXX found a significant effect of sequence with all comparisons were significant ($t = 2.26, df = 13, p < .05$ for the smallest difference). EHS for the first item in the sequence XTXXT was especially elevated. The reason for this is unknown, but we note that it is unique in being a single go stimulus between two no-go stimuli. Mean EHS was shorter for stimuli preceded by two no-go stimuli than for those preceded by a go stimulus ($t = 6.29, df = 13, p < .001$). Though this confounds position in sequence with preceding stimulus it is nonetheless consistent with the last-item effect in suggesting that EHS includes processing for adjacent items. S5 in sequence TXXTT might appear inconsistent as the EHS for S5 is as fast as that of S4 in TTXXT. However, this is due to the last-item effect.

Discussion

Experiment 5 provides several important clues to the components of RT1, which relate to buffering and sequence initiation costs. There was no elevation of Dwell on S1 compared to other location, and thus, no evidence that sequence initiation was delayed. Dwell was longer for target than non-target stimuli as would be expected. When the first target occurred on S1, Dwell was 60 ms longer than when

it occurred on S2, suggestive of a greater difficulty in coordinating the eyes and hands at the outset compared to once the eyes began.

RT1 was approximately 100 slower when on S1 than on S2. As noted above, RT1 on S2 (796 ms) was slower than the comparable response to S4 in TXXTT (668 ms), evidence of an additional first-response cost even after the sequence had begun. It is clearly not the initiation of a sequence that produces the first-response cost. RT1 when on S2 cannot simply reflect delays due to resource conflicts between the transition to the next item and stimulus processing on the current item. Both cases above included a saccade to the next item, with no difference in saccade latency; mean Dwell on S2 when it was the first response was 405 ms compared to 404 ms for S4 in TXXTT. Either the first response is being voluntarily deferred, or there is significant first-time overhead in integrating manual responses into the transition sequence.

Costs for joint eye and hand responses can be seen by comparing the two control conditions, RO (Respond-Only) and RTS (Respond-then-Saccade). RT1 was higher (646 ms) when participants were instructed simply to move their eyes than when told to just respond to the first item (576 ms). Again, however, this cost cannot be attributed to immediate resource conflicts. For example, RT1 in TXXXX (RTS) was 646 ms compared to 890 ms for TXXTT despite the fact that in the latter case the first three positions are identical and the response was made prior to fixating S4. The initial response in both cases was made in exactly the same context. Instead of being a simple response to resource conflicts, coordinating a sequence of manual responses and saccades appears to be determined by expectations for the sequence, not simply the immediate context. Expecting a coordinated sequence of fixations and responses altered the way participants approached the task, especially the first response. This does not reduce the role of resource conflicts, as predicted by central bottleneck model, or other architectures, but suggests that strategies are an integral part of sequence execution.

Does the use of strategies invalidate RO and RTS as appropriate controls? Participants in the RS condition participants could have treated the sequence as a dual-task experiment, completing S1 processing before beginning the eye movement sequence. Indeed, Dwell on S1 in RTS was 534 ms compared to a mean of 460 ms for sequences with a target at S1. Nonetheless, RT1 in RO and RTS corresponds closely to EHS in experimental sequences. When targets were preceded by two no-go stimuli (i.e., XXT), fixation on those stimuli began after the response to the last target. That is, those items were responded to essentially in isolation from previous responses, so EHS would be equivalent to RT. For TTXXT and XTXXT, the XXT item was the final item, so like the RO control there were no subsequent items to fixate. Mean EHS for those two stimuli is 546, compared to RT1 of 576 in the RO control ($t = .96$, $p < .4$). For sequence TXXTT the XXT item occurs in position four with a subsequent fixation and response necessary. Mean EHS was 668 compared with an RT1 of 647 for the RTS control ($t = .63$, $p < .56$). Thus, it is the need to fixate and respond to subsequent items substantially increases the EHS of the current item.

General Discussion

In five experiments, factors affecting the complexity of the sequence were varied to determine whether the previously observed elevation of RT1 resulted from costs in initiating a sequence, or strategies for scheduling the first response. As commonly reported in eye-hand experiments, the eyes were fixated one or two

characters ahead of the character being responded to. It was also clear that this look ahead provided the opportunity for overlap in the processing of adjacent items. In all experiments first response (RT1) in executing a sequence of responses was elevated compared to the mean IRI for subsequent items, replicating earlier eye movement studies using this paradigm (Remington, 2006; S.-C. Wu, & Remington, R. W., 2004; S.-C. Wu, Remington, R. W., & Pashler, H., 2004; S.-C. Wu, Remington, R.W., & Pashler, H., 2007; S.-C. Wu, Remington, R.W., Lewis, R., 2007). RT1 was also elevated compared to conditions where participants responded to only the first item in the sequence (Experiment 5, T____), or responded to the first item then fixated the remaining items without responding (Experiment 5, TXXXX). In the following sections we discuss the evidence for strategic deferral of RT1 and the role of architecture versus strategy in sequence execution.

Components of RT1

Evidence for strategic deferral of RT1 comes primarily from the failure to modulate RT1 with factors that should have affected S1 processing or set-up costs for sequence initiation. RT1 elevation could be explained as lack of preparation, the need to retrieve stimulus-response mappings, or attributable to other factors that tend to slow responses to the first item of a block in discrete-trial experiments (e.g., Altmann, 2007). Experiment 2 found substantial RT1 slowing even after a 1.5 sec preview of S1 (the first item). It would be a stretch to argue that S1 processing was still inefficient after prolonged preview. Also, costs associated with inefficient first-item processing should have affected the two control conditions in Experiment 5, T____ and TXXXX. RT in those conditions was significantly faster than standard sequences. Further, EHS for RT2 (S4) in TXXTT was significantly faster than for any RT1 (S1 or S2), even though the S4 response was well separated from prior items so that its EHS should be equivalent to a standard RT. Finally, comparing RT1 in the TXXXX and TXXTT sequences of Experiment 5 shows clearly that the response to the first item in a sequence was not a function of the local context, but reflected subjects' expectations about the sequence, including items that had yet to be encountered. Indeed, there was no evidence to suggest that S1 processing per se played any role in RT1 elevation.

Likewise, there was no consistent pattern suggesting that set-up costs in sequence initiation delayed RT1. The failure of Experiments 1-3 to find effects of the extent of the display (Experiment 1), the number of items in the display (Experiments 3a & 3b), or the regularity of spacing of items in the display (Experiment 4) on either manual responses or saccades clearly indicates that set-up costs in sequence initiation are not related to complexity or uncertainty in the arrangement of items as would be expected. There is evidence in first-item dwell for minimal set-up costs in sequence initiation. Across experiments there was a slight elevation of Dwell on S1 of 50-80 ms, but that elevation also includes the time to perceive and interpret the display changes that signal the beginning of a trial. This small elevation of Dwell is not sufficient to account for the prolonged RT1 seen in all experiments, and could not have played a role in RT1 on S2.

Strategies in Sequence Execution

Other evidence suggests that subjects adopted a strategy of voluntarily deferring R1. In Experiment 2, for example, RT1 was still 800 ms despite 1.5 seconds of preview. Voluntary postponement of R1 also provides a straightforward explanation of the observed decrease in EHS over the first three items seen across

experiments. This pattern over items is clear in Experiment 3 where it is possible to discern an initial transition phase extending over the first two items, followed by a steady-state phase. It is the pattern that would be expected if the delay in RT1 did not reflect resource conflicts, but a voluntary delay to allow one or more subsequent items to be partially processed before beginning the response sequence.

Additional evidence for voluntary deferment of R1 comes from previous studies in this paradigm, which examined factors affecting selected processing stages (Remington, 2006; S.-C. Wu, Remington, R. W., & Pashler, H., 2004; S.-C. Wu, Remington, R.W., & Pashler, H., 2007; S.-C. Wu, Remington, R.W., Lewis, R., 2007). Manipulations of luminance and stimulus-response compatibility produced an effect on RT1 more than twice that for IRI. For stimulus-response compatibility (see e.g., S.-C. Wu, & Remington, R. W., 2004) in the difficult mapping IRI increased by 150 ms, Dwell by 120 ms, but RT1 increased by over 400 ms. This doubling is approximately what would be expected if R1 were postponed until response selection had been completed on S2, since the increased RS difficulty for both S1 and S2 would have contributed.

If RT1 is being delayed by a deliberate choice to buffer the first response, then what might drive this strategy? One possibility is that withholding the initial response allows the eyes to get ahead, creating the conditions for overlap that make for fluid sequence execution. Buffering the first response (or two) could also protect a regularly timed sequence of responses fluctuations in input timing. Such considerations imply that people approach sequences with the goal of establishing a coupled input-output sequence characterized by a constant rhythm of eye movements and manual responses. Buffering one or more items sets the conditions for this regular, rhythmic sequence. In this regard it is worth noting the close correspondence between IRI and Dwell that emerges in the steady-state phase of execution. This can be seen by comparing Dwell and IRI for S3 – S8 in Experiment 3, which omits the transition phase seen in S1 and S2 as well as the final item effect on S9. The values of Dwell and IRI are similarly close in comparable conditions of previous experiments (Remington, 2006; S.-C. Wu, Remington, R. W., & Pashler, H., 2004; S.-C. Wu, Remington, R.W., & Pashler, H., 2007; S.-C. Wu, Remington, R.W., Lewis, R., 2007). The choice for participants then is when to begin the response sequence. Further investigation may reveal a trade-off between memory cost and efficiency as seen elsewhere (W. D. Gray et al., 2006; Hayhoe, Jovancevic, & Sullivan, 2006; Hayhoe & Land, 1999).

Strategic Resource Scheduling

If buffering the first response is a strategic choice then what role do resource constraints play in determining eye movements and manual responses in a sequence? It is possible, for example, that buffering is a way to remove resource constraints by providing enough slack in the schedule that they never determine performance directly. Framed differently, the question is whether the timing of events in a sequence tells us anything about the underlying resource architecture. It is premature to attempt a definitive answer to this question. What can be done is to see whether resource constraints provide a reasonable account of the eye and hand events in the steady-state phase of the sequence. Below we present two simple models how resources might be scheduled in combination with an initial first-item delay. Both demonstrate in principle how the results could arise from an initial choice of deferring R1 with subsequent responses constrained by the underlying bottleneck structure of resources. They are not intended to be complete detailed models of sequence

execution. Rather, our goal is to explore the consequences of two different ways of scheduling resources.

Model description. The two models are diagrammed in the top-left and top-right panels of Figure 9. Below each is a graph comparing simulated results of each respective model with data from Experiment 3b. As described below the parameters of the model were not estimated from that experiment, but represent average values across all experiments. We simulated the 9-item sequence of Experiment 3b as it contains more of the steady-state component, and the RT1 elevation is closer in value to the average across experiments than is Experiment 3a.

The two models differ chiefly in how central processing for the transition to the next item is scheduled with respect to that of the current stimulus. The top-left diagram shows a simple hypothesized schedule of operations, derived from single central bottleneck theory, in which central processing (RS stage) of each stimulus is completed prior to shifting attention (and the eyes) to the next item. We refer to this as the CP-first model. In the CP-First model only the motor stage is delayed. The top-right diagram of Figure 9 shows a model much, if not all, central processing is done after the transition to the next item. We refer to this as the T-First model. The dark grey boxes represent the bottleneck processes for stimulus processing on each item, the light grey the central bottlenecks associated with the transition from one item to the next.

As our interest is in the effects of constraints, not in assigned functional roles refer to the stages as P = perceptual, C = central, and M = motor, rather than the functional labels of stimulus encoding, response selection, and response execution. We let T = the central demands of item transition. In a departure from strict single channel central bottleneck theory, each model posits some constraint on the execution of Motor operations. In the CP-First model, Motor operations occur at the end of the Central processing for the transition, not the end of Central processing for the task (which would typically be Response Selection). This is equivalent to asserting a motor initiation bottleneck (not explicitly represented in Figure 9), or at least constraints on the parallel execution of responses. An analogous assumption is incorporated into the T-First model, where Motor operations for one stimulus cannot be initiated in parallel with ongoing Central operations on another.

To fix model parameters we estimated the range of times to perform the sum of the Perceptual, Cognitive, and Motor operations of stimulus processing to be 550-600 ms based from RT in the Respond-Only condition of Experiment 5. This range corresponds very well with the EHS for final items preceded by no-go stimuli in Experiment 5, as well as the EHS for the final item in Experiments 2, 3b, and 4 (final-item EHS is closer to 700 ms for Experiment 1, and Experiment 3a, possibly because IRI and Dwell are also higher). As argued earlier, EHS for the final item should reflect only stimulus processing, as no transition needs to be scheduled. The average final-item reduction if EHS across experiments was close to 100 ms, which we used as the estimate of T, the central processing for the transition. With this assumption, the model also dictates that the final IRI should reflect only the C and M stages plus the saccade duration. Across experiments the average difference between the final EHS and the final IRI is approximately 150 - 200 ms, an estimate of the duration of perceptual processing. Varying the times of Perceptual, Central, and Motor stages from 150 - 200 ms, with the constraint that the total be between 550 - 600 ms, made no significant effect on the simulation results. Here we present data for the CP-First model using 200 ms for each stage. For the T-First model values of P, C, and M were 300, 150, and 150 ms, respectively. A 60 ms delay was added to the T stage of S1 in

both models to reflect the average saccade delay across experiments. Various durations of voluntary deferment were tested with the most accurate being 200 ms, which was the value used here.

Simulations & Results. Monte-Carlo simulations were conducted for each model using the freely available R statistical package (R 2.4.1 ©2004 – 2006). Each simulation consisted of 1000 trials. Mean values were assigned to the P, C, M, and T operations as described above, as well as a 30 ms mean saccade duration. The standard deviation for each parameter was set to 50% of mean value. Initial simulations found no meaningful differences between setting the standard deviation to 25% or 50% of the mean value. For each parameter, 1000 values were drawn, one for each “trial”, from a normal distribution with the assigned mean and standard deviation for that parameter. Predicted times for each saccade and manual response were derived by applying the rules specified by each model.

The graphs below each diagram show that both models were able to capture main features of the Dwell, EHS, and IRI of Experiment 3b. RT1 itself is not of importance since its value was estimated directly from the data of Experiment 5. However, the consequence of deferring RT1 can be seen in the shortened IRI at S2, which was observed across experiments, though EHS for S2 is lower than observed. The simulations also capture the rise in IRI from S2. Although IRI is statistically unaffected by stimulus position, from S2 to the penultimate stimulus there is a gradual increase. In the CP-First model this occurs as a step increase from S2 to S3, but its gradual nature is somewhat better captured by T-First. Overall, T-First better replicates the pattern of data. We emphasize again that model parameters were not chosen from this experiment but from rough averages across all experiments. The intent was not to fit the data from any one study but to see whether resource constraints in combination with voluntary RI deferment could reproduce the pattern seen across experiments. For this reason, the models themselves were kept simple. Yet, despite the simplicity and crude parameter assignment, simulation outcomes correspond well to the data.

Model implications. Aside from the ability of these simple models to reproduce the pattern of results from a wide range of experiments, two features of them deserve further comment. First, the difference between the CP-First and T-First models relates to the ongoing debate on whether eye movements can occur after perceptual processing (e.g., Salthouse & Ellis, 1980; Sanders & van Duren, 1998) or are executed only after some central processing has been completed (see e.g., Inhoff, Briihl, Bohemier, & Wang, 1992; Inhoff & Gordon, 1997). The bulk of the evidence favors the latter. For example, as reported above, we have found that eye movement latencies are increased with increases in central processing difficulty, and studies of reading have shown that increases in lexical and semantic difficulty increase fixation duration. Yet here, the T-First model performed somewhat better than the CP-First model. The principle reason for this is that T-First decouples the saccade (transition to the next item) from the stimulus processing. This decoupling seems to be important in capturing the decline in EHS and the subtle increase in IRI up to the penultimate stimulus. In an earlier model (Remington, 2006) we found very good fits with a model in which the individual saccades were initiated at regular intervals without reference to the state of stimulus processing. The model was similar to other non-process models in that eye movement timing was based on a global estimate of processing duration (see I. T. C. Hooge & Erkelens, 1998; L. T. C. Hooge & Erkelens, 1996; Legge, Klitz, & Tjan, 1997). More work will be required to see how

the simple T-First model can be modified to reflect the effect of processing difficulty on eye fixation durations.

Secondly, both models introduce constraints on the initiation of Motor execution. While not a feature of classic central bottleneck theory, recent work has found some support for bottlenecks associated with motor initiation or execution (McLeod & Hume, 1994; Meyer & Kieras, 1997a, 1997b). Several cognitive architectures assume that each motor act is initiated by a brief central operation (Card, Moran, & Newell, 1983; John, 1996; John et al., 2002; Vera et al., 2005). Recent findings (Bratzke et al., 2008) support a motor initiation bottleneck by showing that more demanding response execution produces slack associated with a bottleneck process. The introduction of a motor initiation bottleneck raises the issue of whether it is necessary to retain the assumption of a central bottleneck, which has been previously challenged (Meyer & Kieras, 1997a, 1997b). Future research exploring resource assumptions other than those of central bottleneck theory will help resolve this issue, and provide a richer understanding of the role of strategy and resource constraint in performing common daily tasks.

Conclusions

We showed that the timing of manual responses in a sequence is characterized by a strategic deferment of the response to the first item. This is further evidence of "soft" constraints that emerge when actions are done in the context of other actions rather than in isolated discrete-trial presentation. Nonetheless, the patterns observed were very regular across experiments and participants, suggesting that the choice of strategies was not arbitrary, but followed well-defined rules. The results of two models of resource scheduling in sequence execution showed that in principle it is possible to derive predictions of the overt eye movement and manual responses by combining the assumption of voluntary deferment with an underlying resource model.

References

- Altmann, E. M. (2007). Comparing switch costs: Alternating runs and explicit cuing. *Journal of Experimental Psychology-Learning Memory and Cognition*, 33(3), 475-483.
- Bratzke, D., Ulrich, R., Rolke, B., Schroter, H., Jentzsch, I., & Leuthold, H. (2008). Motor limitation in dual-task processing with different effectors. *Quarterly Journal of Experimental Psychology*, 61(9), 1385-1399.
- Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review*, 108(4), 847-869.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, N.J: L. Erlbaum Associates.
- Dessouky, M. I., Moray, N., & Kijowski, B. (1995). Taxonomy of scheduling systems as a basis for the study of strategic behavior. *Human Factors*, 37(3), 443-472.
- Epelboim, J., Collewyn, H., Edwards, M. A., Erkelens, C. J., Kowler, E., Pizlo, Z., et al. (1993). Coordination of Eyes Head and Hand in a Natural 3-D Tapping Task. *Investigative Ophthalmology & Visual Science*, 34(4), 1502-1502.
- Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research*, 41(12), 1561-1574.
- Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 266(1436), 2435-2440.
- Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS Analysis for Predicting and Explaining Real-World Task Performance. *Human-Computer Interaction*, 8(3), 237 - 309.
- Gray, W. D., Sims, C. R., Fu, W. T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review*, 113(3), 461-482.
- Hayhoe, M., Jovancevic, J., & Sullivan, B. (2006). Control of attention and gaze in complex environments. *Australian Journal of Psychology*, 58, 78-79.
- Hayhoe, M., & Land, M. (1999). Coordination of eye and hand movements in a normal visual environment. *Investigative Ophthalmology & Visual Science*, 40(4), S380-S380.
- Hooge, I. T. C., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, 38(9), 1295-1302.
- Hooge, L. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, 58(7), 969-976.

- Inhoff, A. W., Briihl, D., Bohemier, G., & Wang, J. (1992). Eye-Hand Span and Coding of Text during Copytyping. *Journal of Experimental Psychology-Learning Memory and Cognition*, 18(2), 298-306.
- Inhoff, A. W., & Gordon, A. M. (1997). Eye movements and eye-hand coordination during typing. *Current Directions in Psychological Science*, 6(6), 153-157.
- John, B. E. (1996). TYPIST: A theory of performance in skilled typing. *Human-Computer Interaction*, 11(4), 321-355.
- John, B. E., Vera, A., Matessa, M., Freed, M., & Remington, R. (2002). Automating CPM-GOMS. In *Proceedings of CHI, April 20-25*. Minneapolis, MN: ACM.
- Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12(4), 391-438.
- Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr Chips: An ideal-observer model of reading. *Psychological Review*, 104(3), 524-553.
- Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology-Human Perception and Performance*, 29(3), 575-599.
- McLeod, P., & Hume, M. (1994). Overlapping Mental Operations in Serial Performance with Preview - Typing - a Reply. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 47(1), 193-199.
- Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance .1. Basic mechanisms. *Psychological Review*, 104(1), 3-65.
- Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance .2. Accounts of psychological refractory-period phenomena. *Psychological Review*, 104(4), 749-791.
- Pashler, H. (1984). Processing Stages in Overlapping Tasks - Evidence for a Central Bottleneck. *Journal of Experimental Psychology-Human Perception and Performance*, 10(3), 358-377.
- Pashler, H. (1994). Overlapping Mental Operations in Serial Performance with Preview. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 47(1), 161-191.
- Pashler, H., Carrier, M., & Hoffman, J. (1993). Saccadic Eye-Movements and Dual-Task Interference. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 46(1), 51-82.
- Pashler, H., & Johnston, J. C. (1989). Chronometric Evidence for Central Postponement in Temporally Overlapping Tasks. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 41(1), 19-45.

- Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, 106(4), 643-675.
- Reichle, E. D., & Laurent, P. A. (2006). Using reinforcement learning to understand the emergence of "intelligent" eye-movement behavior during reading. *Psychological Review*, 113(2), 390-408.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105(1), 125-157.
- Remington, R. W., Lewis, R., & Wu, S-C. (2006). Scheduling mental operations in a multiple-response sequence: Modeling the effects of a strategy to minimize variance in the timing of saccades. In D. Fum, F. del Missier, & A. Stocco (Ed.), *Proceedings of the 7th International Conference on Cognitive Modeling, Trieste, Italy. April 2-6, 2006*.
- Rogers, R. D., & Monsell, S. (1995). Costs of a Predictable Switch between Simple Cognitive Tasks. *Journal of Experimental Psychology-General*, 124(2), 207-231.
- Rosenbaum, D. A. (2002). Motor Control. In H. Pashler & S. Yantis (Eds.), *Stevens' handbook of experimental psychology vol. 1: Sensation and perception, 3rd edition*. (Vol. 3, pp. 315-340). New York: John Wiley & Sons, Inc.
- Ruthruff, E., Johnston, J. C., Van Selst, M., Whitsell, S., & Remington, R. (2003). Vanishing dual-task interference after practice: Has the bottleneck been eliminated or is it merely latent? *Journal of Experimental Psychology-Human Perception and Performance*, 29(2), 280-289.
- Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology-Human Perception and Performance*, 27(6), 1404-1419.
- Salthouse, T. A. (1986). Perceptual, Cognitive, and Motoric Aspects of Transcription Typing. *Psychological Bulletin*, 99(3), 303-319.
- Salthouse, T. A., & Ellis, C. L. (1980). Determinants of Eye-Fixation Duration. *American Journal of Psychology*, 93(2), 207-234.
- Sanders, A. F., & van Duren, L. L. (1998). Stimulus control of visual fixation duration in a single saccade paradigm. *Acta Psychologica*, 99(2), 163-176.
- Vera, A. H., John, B. E., Remington, R., Matessa, M., & Freed, M. A. (2005). Automating human-performance modeling at the millisecond level. *Human-Computer Interaction*, 20(3), 225-265.
- Verwey, W. B. (2003). Effect of sequence length on the execution of familiar keying sequences: Lasting segmentation and preparation? *Journal of Motor Behavior*, 35(4), 343-354.
- Welford, A. T. (1952). The 'psychological refractory period' and the timing of high-speed performance -- a review and a theory. *British Journal of Psychology*, 43, 2-19.

Wu, S.-C., & Remington, R. W. (2004). Coordination of component mental operations in a multiple-response task. In S. N. Spencer (Ed.), *Proceedings of the Eye Tracking Research and Applications Symposium 2004*. New York: ACM SIGGRAPH.

Wu, S.-C., Remington, R. W., & Pashler, H. (2004). Coordination of component mental operations in sequences of discrete responses. In K. Forbus, D. Gentner, & T. Regier (Ed.), *Proceeding of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 1476-1481). Mahwah, NJ: LEA.

Wu, S.-C., Remington, R.W., & Pashler, H. (2007). Scheduling of Eye Movements and Manual Responses in Performing a Sequence of Choice Responses: Empirical Data and Model. In R. L. Lewis, T.A. Polk, & J.E. Laird (Ed.), *Proceedings of the 8th International Conference on Cognitive Modeling*. Ann Arbor, MI July 26-29, 2007.

Wu, S.-C., Remington, R.W., Lewis, R. (2007). Modeling the scheduling of eye movements and manual responses in performing a sequence of discrete trials. In R. Sun (Ed.), *Proceedings of the 28th Annual Cognitive Society, Vancouver, BC, Canada, July 26-29*.

Table 1

Experiment	Cond	RT1	F2	F3	RT1-F2	RT1-F3
1	Wide	1131	643	1259	488	-129
2	Preview	800	417	814	383	-14
2	No Preview	953	498	921	455	32
3a	3	1401	711	1431	690	-30
3a	5	1408	700	1421	708	-14
3a	9	1429	700	1433	729	-4
3b	3	1076	497	930	579	146
3b	5	1114	493	924	621	190
3b	9	1156	526	972	630	183
4	Regular	962	510	930	452	32
4	Irregular	938	517	931	421	7
5	S1=T	890	519	1048	371	-158
Mean		1105	561	1085	544	20

Table 1 indicates when RT1 occurred relative to S2 processing for each condition in each experiment (for Experiment 5, only sequences beginning with a target are included). F2 = the time in the trial at which S2 fixation began. F3 = the time S3 fixation began. F3-RT1 indicates when RT1 occurred relative to beginning fixation on S3. Negative values mean that RT1 occurred after fixating on S3. RT1-F2 indicates the amount of time for S2 processing available prior to RT1. Values of F3-RT1 and RT1-F2 were computed for each participant and averaged.

FIGURE 1

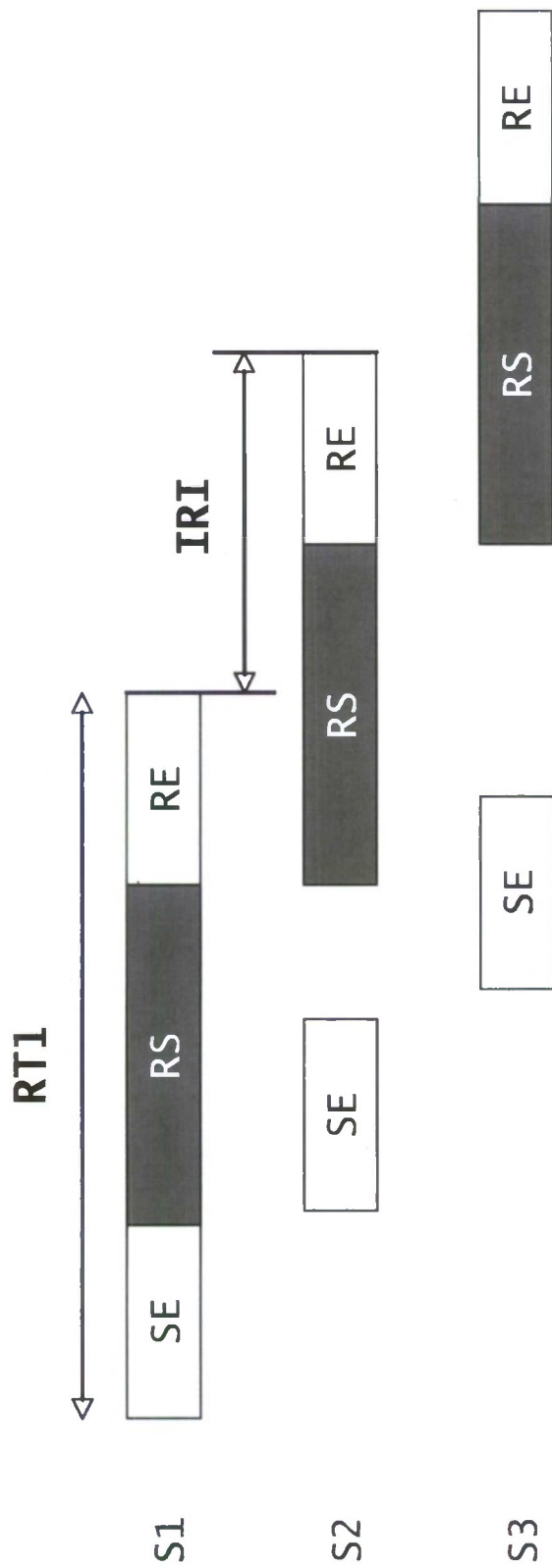


FIGURE 2

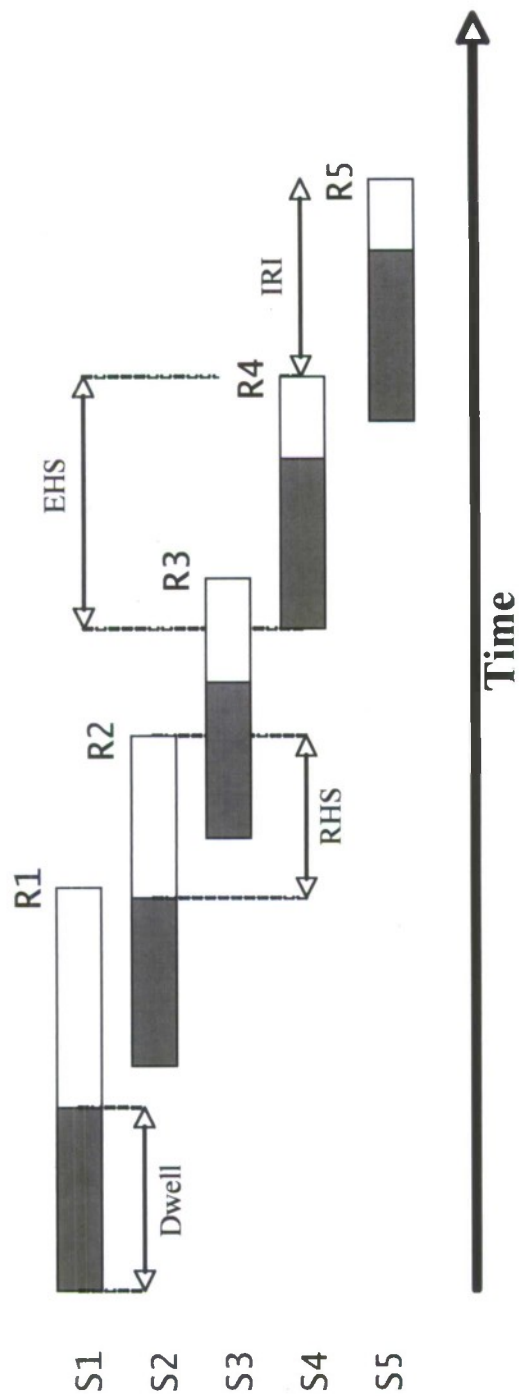


Figure 3

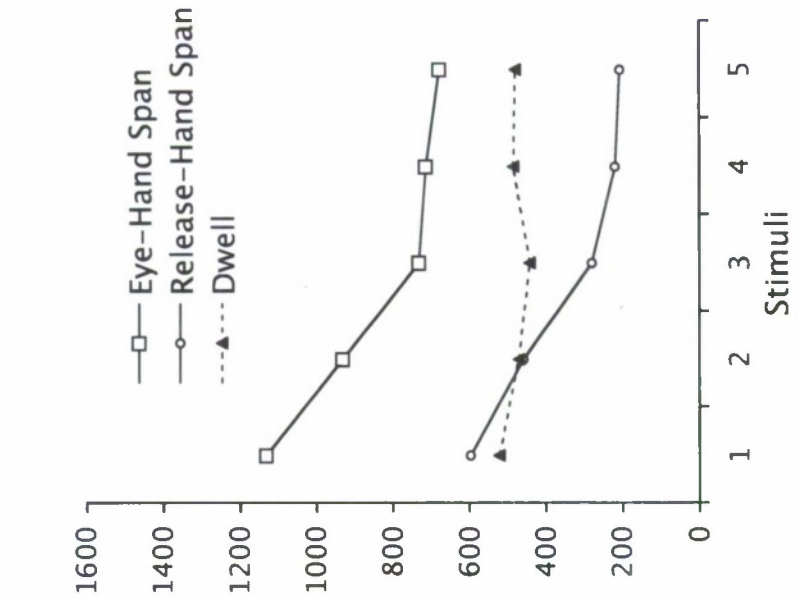
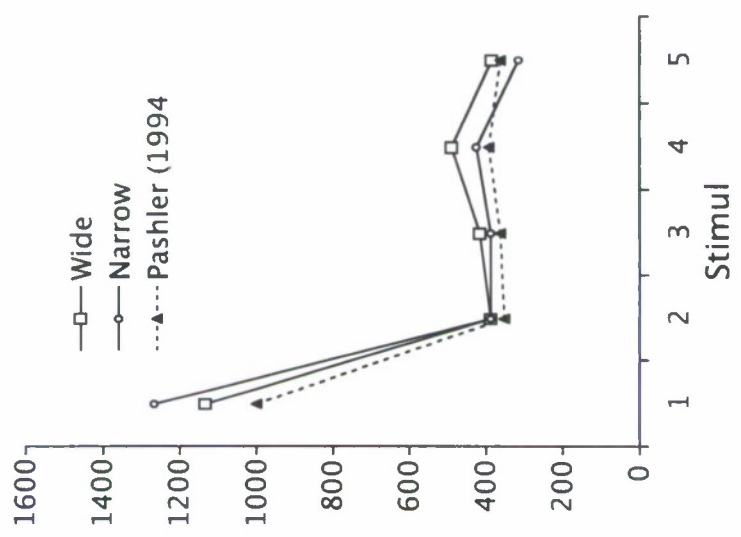


Figure 4

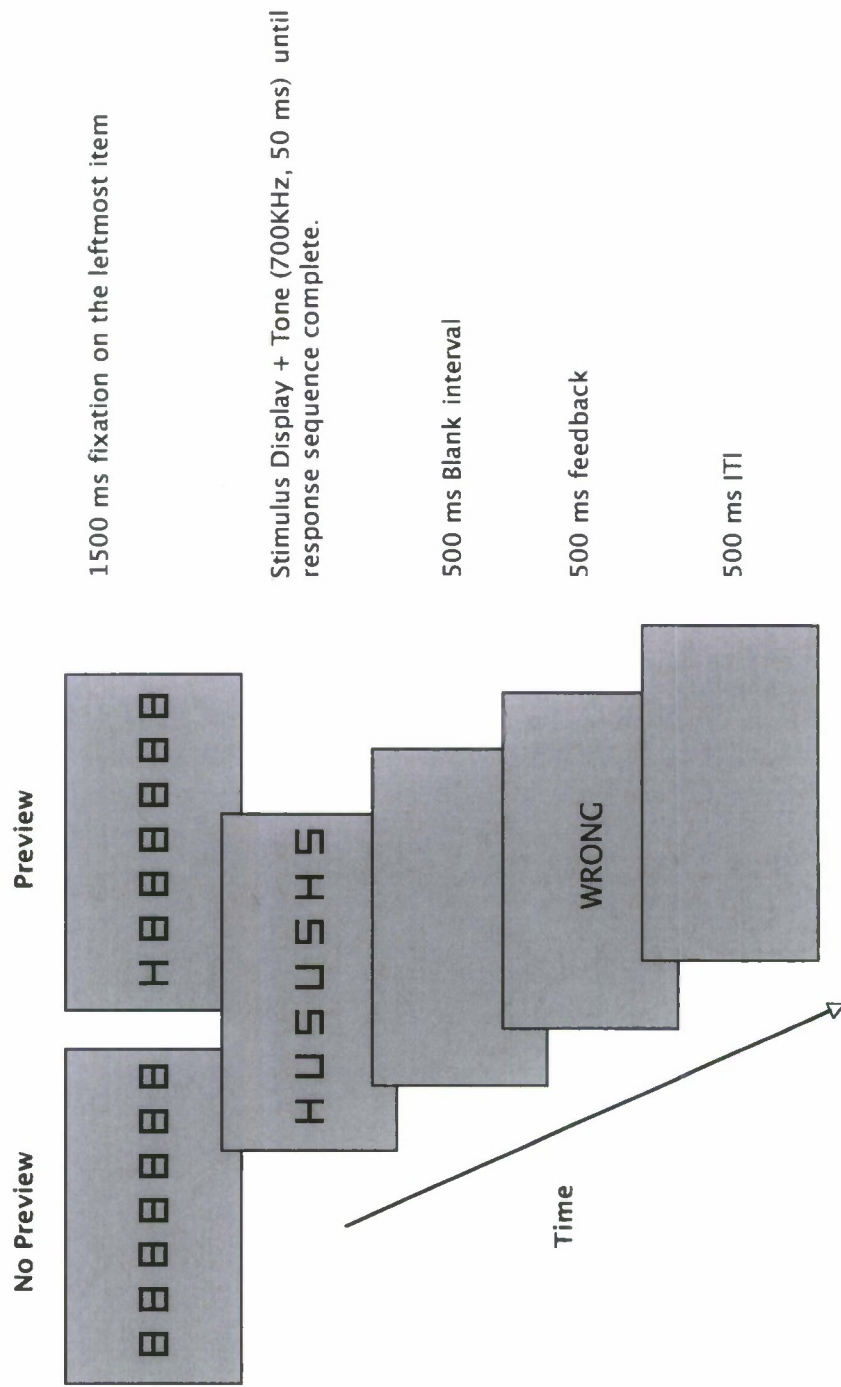


Figure 5

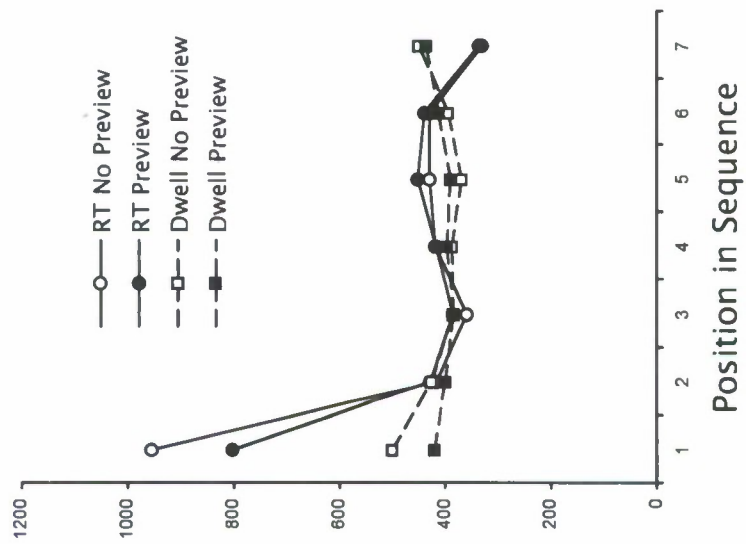


Figure 6

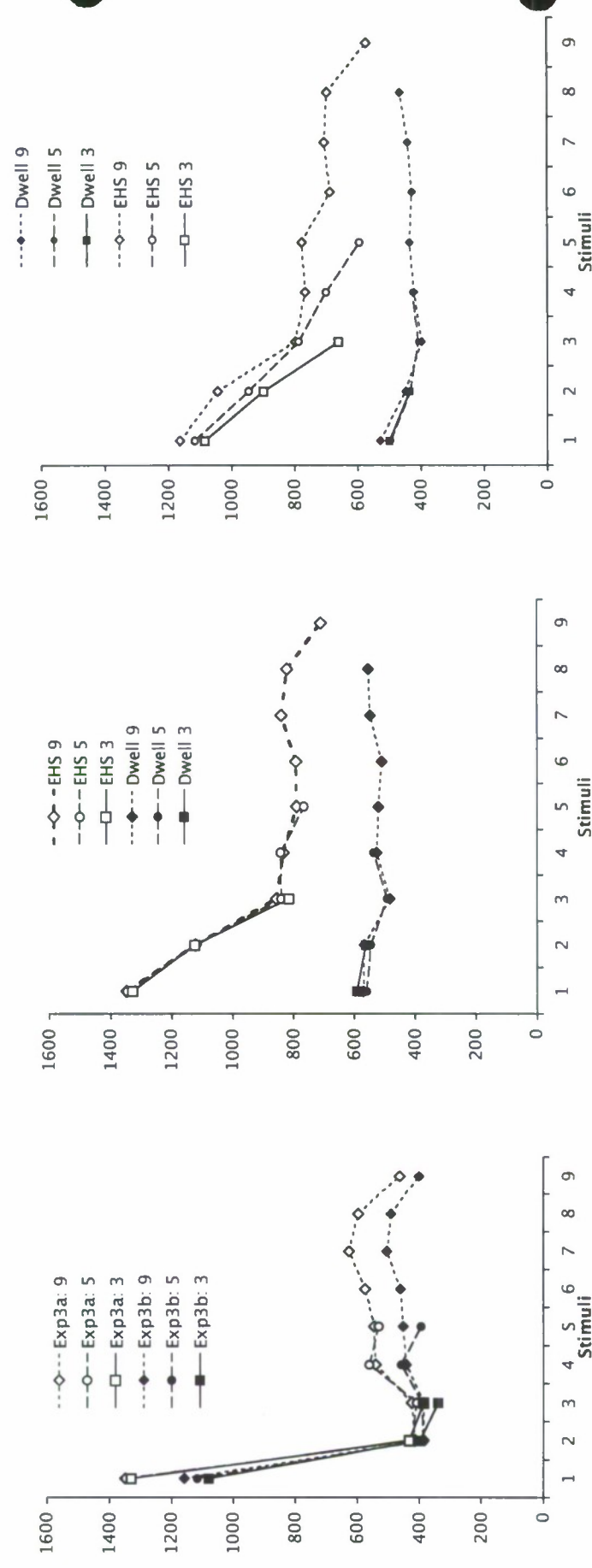


Figure 7

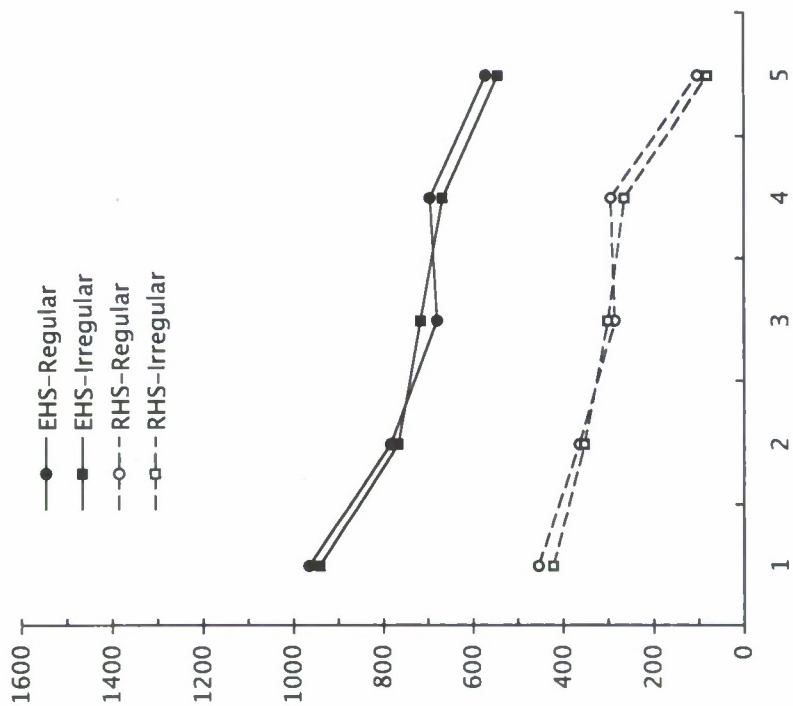
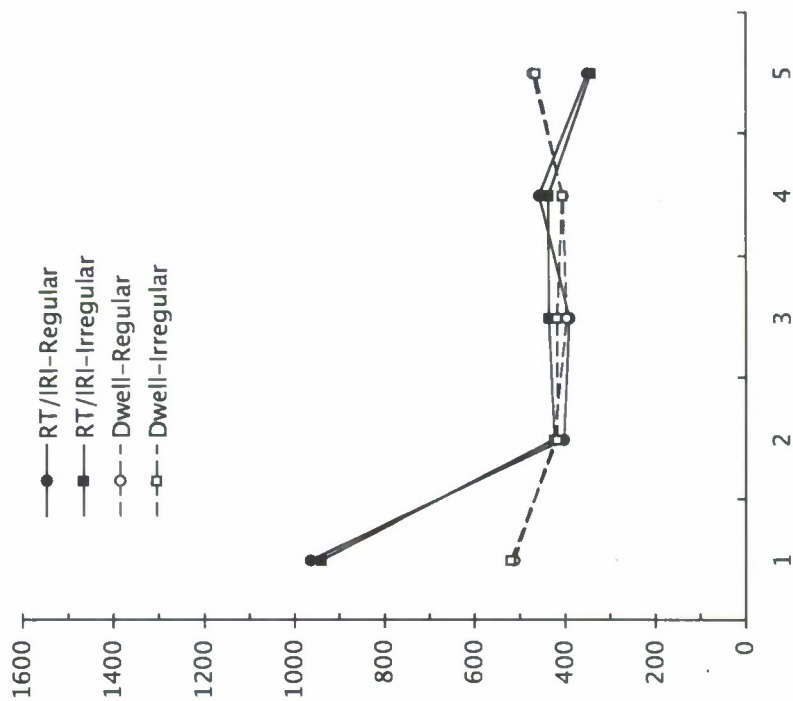


Figure 8

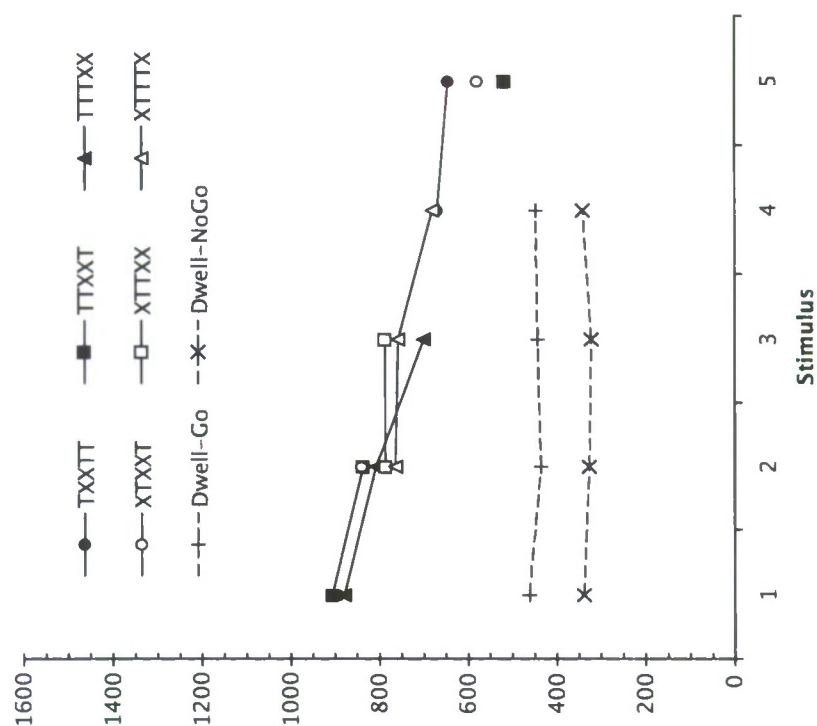
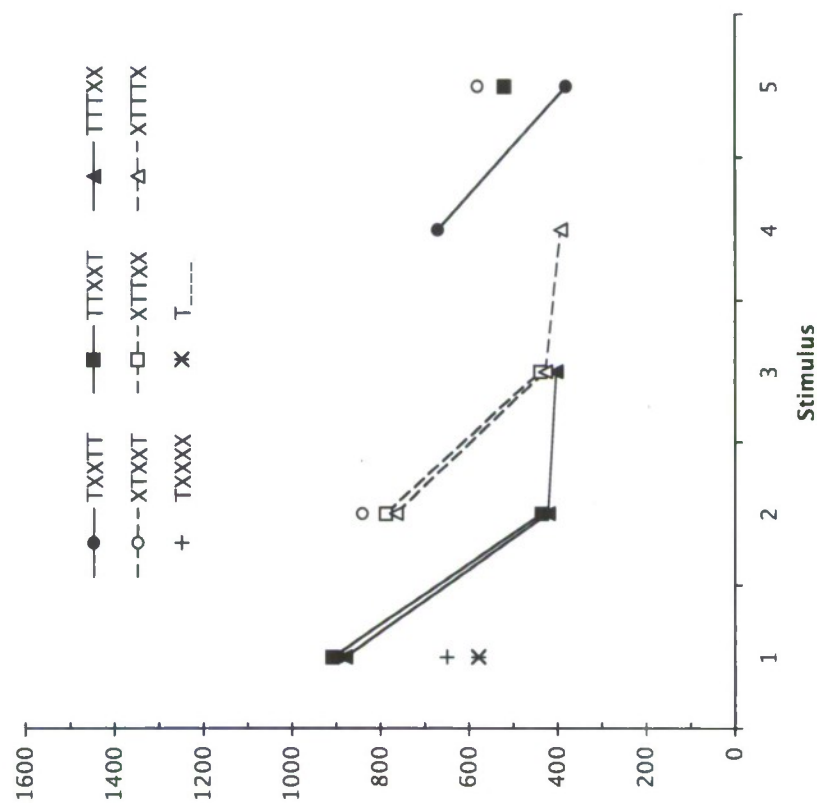
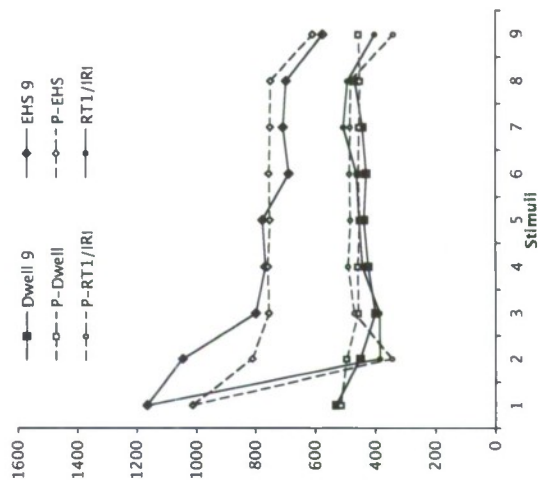
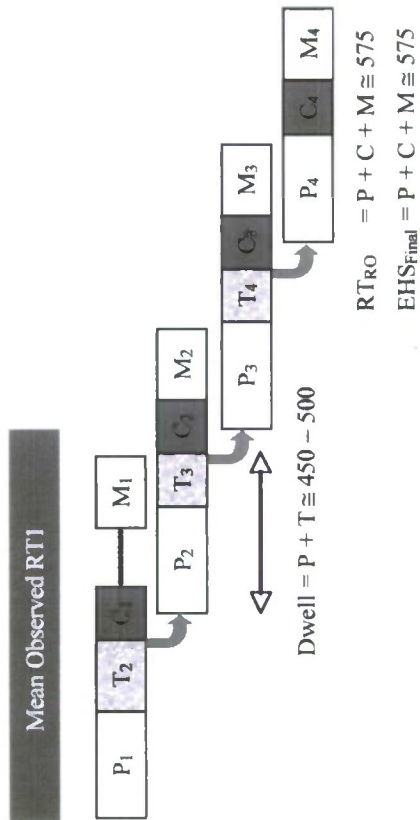
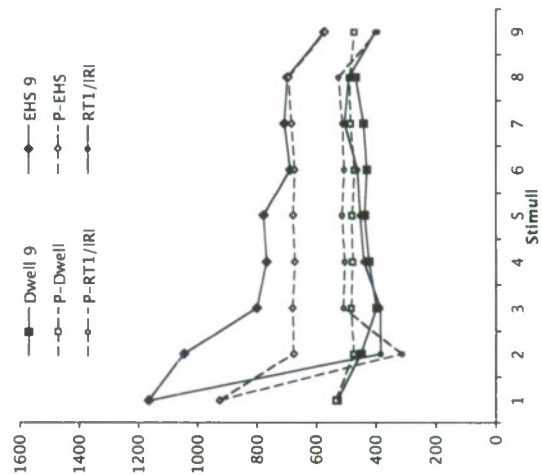
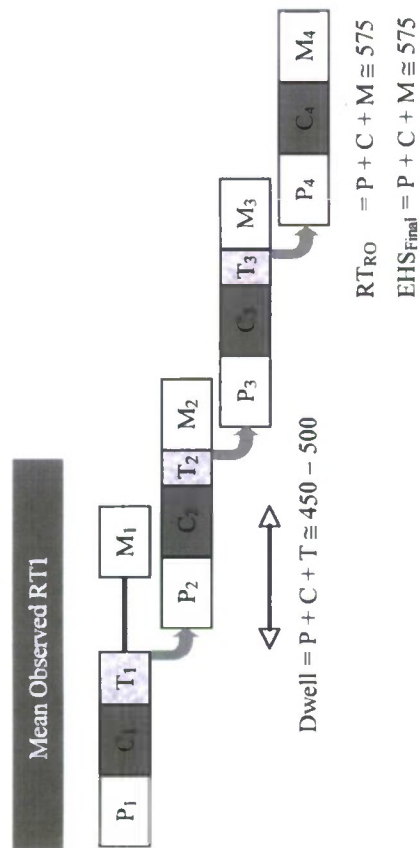


Figure 9



Appendix B: Paper on contingencies in eye movement scheduling to be submitted to
Psychological Science

What is Overlapped in Coordinated Eye-Hand Sequences

Roger Remington
University of Queensland

Shu-Chieh Wu
San Jose State University and NASA Ames Research Center

Harold Pashler
University of California at San Diego

Many common activities are done using a well-learned sequence of actions, with tight coordination of motor responses and eye movements. Studies of activities such as typing (Inhoff & Gordon, 1997; Inhoff & Jian, 1992; Salthouse, 1986), reading (Rayner & Pollatsek, 1981), golf (Vickers, 1992), driving (Land & LEE, 1994) as well as other eye-hand tasks (Land & Hayhoe, 2001; Land & Mcleod, 2000; Pelz, Hayhoe, & Loeber, 2001) consistently show that people look ahead, making an eye movement to the stimulus for the next action before responding to the current one. The general view is that looking ahead takes advantage of parallelism in the human cognitive architecture, allowing portions of the stimulus processing for successive actions to overlap to create a smooth, fluid sequence. Indeed, laboratory paradigms often force overlap in the mental processing of two or more stimuli to reveal parallel and serial components of the underlying architecture (Pashler, 1984; Welford, 1952). The question here is when people fixate ahead of the response in executing an action sequence what processing do they naturally overlap.

An answer to this question must address how the initiation of the saccade is related to the processing of stimuli in the sequence. It has proven useful to decompose simple choice RT tasks into a series of functional stages (Pashler, 1984; Sternberg, 1969; Welford, 1952) consisting of Stimulus Encoding (SE), Response Selection (RS), and Response Execution (RE) in that order. According to Central Bottleneck Theory (Pashler, 1984; Sternberg, 1969; Welford, 1952) SE and RE are done by specialized systems and can execute in parallel. RS requires generalized central processing resources so that only RS for one stimulus can execute at a time. For linguistic tasks this functional diagram may be elaborated to include lexical access and semantic processing (see e.g., Reichle, Pollatsek, & Rayner, 2006), with additional assumptions about the central demands of those functional stages.

In theory, the saccade could be made sometime during or after stimulus encoding; continued fixation plays at best a marginal role once the stimulus has been perceived. In some cases the saccade does appear to be made on completion of perceptual processing (Salthouse & Ellis, 1980; Salthouse, Ellis, Diener, & Somberg, 1981; Sanders & van Duren, 1998). Such does not appear to be the case for reading (Rayner & Pollatsek, 1981; Reichle et al., 2006) or transcript typing (Inhoff & Gordon, 1997; Inhoff & Jian, 1992; Salthouse, 1986), however, as increases in fixation duration (dwell) on a word are influenced by the difficulty of post-perceptual factors, such as word frequency or motor difficulty (e.g., Inhoff, Rosenbaum, Gordon,

& Campbell, 1984), suggesting a later locus. We (S.-C. Wu, & Remington, R. W., 2004; S.-C. Wu, Remington, R. W., & Pashler, H., 2004) have shown that incompatible stimulus-response mappings, presumably affecting RS, increase dwell time, suggesting that the saccade is not initiated until at least some RS processing is complete.

Saccade initiation need not be linked directly to the completion of a processing stage. Effects of difficulty on dwell can also be accounted for by assuming that the saccade is timed to provide new input when needed. For example, saccades could be timed so that SE on N+1 finished just as RS on N finished. That way the central stages of each task could be done without waiting, maximizing the benefits of overlap (e.g., Reichle & Laurent, 2006; Remington, 2006). Optimal models information access have been successfully applied to memory (Anderson), library search (Pirolli, Card, Chi), and increasingly to reading (Legge, Hooven, Klitz, Mansfield, & Tjan, 2002; Legge, Klitz, & Tjan, 1997; Reichle & Laurent, 2006; Reichle et al., 2006). Figure 1 depicts various options for scheduling saccades.

The interpretation of saccade timing in reading and typing is complicated by the syntax and semantics of language, and, for reading, by the role of comprehension. Comprehension lacks a clear online marker making it difficult to determine when the saccade is initiated relative to progress toward or completion of comprehension. Keystroke timing in typing can likewise be affected by grammatical and orthographic factors, includes chunking of familiar sequences, and involves motor sequences with varying difficulty (see e.g., John, 1996; Salthouse, 1986). In the experiment reported below we investigated how central processing demands affected saccade initiation using a paradigm adapted from Pashler (1994) that avoids complexities associated with typing and reading. Subjects made a series of speeded choice responses to a linearly arrayed set of 5 letters (see also, Remington, 2006; S.-C. Wu, & Remington, R. W., 2004; S.-C. Wu, Remington, R. W., & Pashler, H., 2004; S.-C. Wu, Remington, R.W., & Pashler, H., 2007; S.-C. Wu, Remington, R.W., Lewis, R., 2007). Each stimulus in a sequence is characterized by an eye-hand span (EHS) that measures the time from first fixating the stimulus to its response. The EHS has two components: Dwell measures the time spent fixating the stimulus, release-hand span (RHS) the time from the saccade to the next stimulus until the response. Manual output is characterized by response time to the first stimulus (RT1) and inter-response intervals (IRI) for successive stimuli.

If saccades are initiated upon the completion of RS then RHS should reflect only response execution time (RE). According to central bottleneck theory, RE can proceed in parallel with stimulus encoding (SE) and with RS. So long as RHS produces no obvious perceptual or motor conflicts (e.g., requiring the eyes to move in different directions), overlapping RE on stimulus N with SE or RS on N+1 should not create a conflict. Moreover, RHS should be unaffected by RS difficulty factors, as they affect processing prior to the initiation of the saccade. Instead, difficulty should be fully reflected in the Dwell time. If, on the other hand, RS is completed after the saccade, then RS difficulty should be reflected in RHS (for a version of this model see, Salthouse & Ellis, 1980; Salthouse et al., 1981). An intermediate model is possible, of course, in which RS is not a single, monolithic stage, but can be decomposed into substages. If the difficulty factor affects a substage prior to saccade initiation then it should be reflected in Dwell; if it affects a later substage then it will be reflected in RHS. Only the two models that posit some or all RS processing completed prior to the saccade predict that Dwell on N+1 would be affected by the difficulty of RS on N. If Dwell depended only on the completion of SE, and not on a bottleneck stage, there would be no mechanism for processing on N to influence Dwell for N+1.

What predictions regarding the affects of difficulty follow from models in which saccade initiation is based on estimated processing times with the goal of optimizing some aspect of performance? Generally speaking two opposing views on estimation have emerged: in one saccade timing is based directly on processing of the immediate stimulus (Rayner, 1998; Rayner & Pollatsek, 1981; Reichle & Laurent, 2006), in the other on global information derived from the recent history or current context (for examples in visual search and reading see Hooze & Erkelens, 1996; Legge et al., 2002; Legge et al., 1997). Both predict increased dwell times with increased central processing requirements. However, global estimation further predicts that the context should influence the timing of the saccade as dwell is weighted function of the current stimulus processing with that of recent stimuli. From this it follows that dwell for a hard stimulus would be underestimated in a easy context, and overestimated for an easy item in a hard context. For a hard stimulus in an easy context the underestimate of dwell would mean a longer RHS than an equivalent stimulus in the context of other hard stimuli. Additionally, if the RHS

contains unfinished central bottleneck processing, the difficulty of stimulus N could also elevate dwell on N+1.

We (Wu, Remington, Pashler, 2004; Wu, Remington, Pashler, 2006) have previously described an experiment in which the difficulty of stimulus-response mapping was varied between blocks. Mean IRI was approximately 120 ms slower for the hard mappings and did not vary from S2 to S5. Dwell time was also elevated by about the same amount. Significantly mean RHS at positions S3 and S4 showed no effect of S-R compatibility. These are the two critical stimulus positions. RHS for the first two positions appear to be elevated due to a strategy of deferring the first response, while RHS for the final stimulus (S5) is not meaningful as there is no further stimulus to fixate. The results support the view that RS is largely completed during the dwell. The critical test of this, and of the estimation accounts, involves the context in which hard and easy items are placed. In Wu, et al. (2004) every trial in a contiguous block of trials was of the same difficulty. Here we vary difficulty within a trial, mixing hard and easy stimuli in a sequence to expose the effects of the global and local context.

Experiment Overview

Subjects made a series of speeded choice responses to 5 stimuli arrayed linearly across a CRT screen. Each stimulus consisted of a 2x2 matrix containing one, two, or three identical digits. The task was to quickly indicate how many digits were present. For compatible stimuli the number of digits was the same as the digit value (e.g., 2, 2). For incompatible stimuli the digit value conflicted with the number (e.g., 3, 3). At issue is whether the extra processing associated with the incompatible stimulus will be fully reflected in the increase in Dwell, or whether RHS will also be extended. The construction of sequences and design also allowed examination of the effects of context. Pure sequences were homogenous with respect to difficulty, all hard or all easy. Mixed sequences included one or two stimuli from the non-dominant condition. The experiment was conducted in two sessions, one predominantly easy, the other hard. The design makes it possible to assess the contributions of the immediate context, defined by the difficulty of preceding stimuli as well as the global context, defined by the session.

Method

Nineteen participants were recruited from the NASA Ames Research Center participant pool. They were all undergraduate students from local universities and

community colleges. All reported normal or corrected to normal vision. They participated for class credit or were paid \$30 plus travel expenses.

Apparatus & Stimuli. A Pentium 4 PC controlled the presentation of responses, collection of responses, and storage of data. A separate Pentium 4 computer controlled eye movement recording. Eye movements were monitored with a head-mounted video-based eye tracking system (Applied Sciences Laboratory, Model 501) sampling at 120Hz with a spatial precision of approximately 0.5° visual angle. Eye position was determined by computing the distance between the center of the pupil and corneal reflection of the left eye. Experiments were carried out in a quiet, well-lit room with participants seated approximately 60 cm from a 21" CRT display with a 70 Hz refresh rate used for stimulus presentation.

The primary stimulus display consisted of a row of five 2x2 matrices centered at the middle of the display. Each matrix subtended 0.34° in height, presented at a luminance of 11.7 cd/m², spaced approximately 5.5° apart. Stimuli were the numbers 1, 2, and 3 presented inside a cell of the matrix. One, two, or three cells of the matrix contained numbers, and participants responded by pressing the V, B, N keys, respectively, to indicate whether there were 1, 2, or 3 numbers present.

Procedure. A total of 384 trials were presented in 4 blocks of 96 trials each, preceded by 24 practice trials. Blocks were designated Easy and Hard to indicate the dominant stimulus type for that block: compatible for Easy, incompatible for Hard. Subjects completed two blocks of one type followed by two block of the other. Block order was counterbalanced across subjects; half did the Easy blocks first, half the Hard. For Easy blocks the 4 sequences were: EEEEE, EHEEE, EEHEE, HEEHE. For the Hard blocks the 4 sequences were the mirror of the Easy: HHHHH, HEHHH, HHEHH, EHHEH.

Each trial began with the presentation of a white fixation cross (0.3°) in the center of the display. After the participant had maintained fixation within a 6° radius around the fixation for 500 ms, the fixation was erased and a small filled square (0.34°) appeared at the leftmost stimulus position. Participants were instructed to fixate the small square and maintain fixation until the stimuli were presented. The small square remained for 1 sec, followed by a blank interval of 500 ms, after which the sequence was presented. Eye movement recording began the moment the small square appeared over the location of the leftmost item, and ended after the participant had responded to the rightmost stimulus. A calibration procedure was administered

before each block of trials to maintain accuracy of recordings. The characters were erased after the participant had responded to the rightmost character. The next trial began following an inter-trial-interval of 250 ms.

Participants were given a written description of the task, which was reviewed with the experimenter. They were instructed to respond to each item as quickly and accurately as they could and not to group their responses.

Manual responses and eye fixations for each item were recorded. Eye fixation samples were analysed offline to classify them into saccades or fixations, and assign fixations to stimuli. Because the stimuli were arrayed horizontally at the same vertical screen position, all analyses were based on horizontal (x-axis) movements only. A saccade was defined as a movement velocity exceeding $30^\circ/\text{s}$ or movement acceleration exceeding $3000^\circ/\text{s}$. A fixation was defined as movement velocity below $30^\circ/\text{s}$ or movement acceleration fell below $-3000^\circ/\text{s}$. A fixation was assigned to the nearest stimulus letter position and its duration was calculated by summing all contiguous individual fixations on a designated target region. Once a fixation on an item ended subsequent fixations on that item were considered regressions. Fixations above or below the stimulus array, or to either side of it, were considered anomalous and omitted in the analyses.

Results & Discussion

Of the 24 participants seven tested 12 had to be excluded because of malfunctions in the eye tracking resulting in lost or corrupted data, and one because the pattern of eye movements was too erratic. Manual and eye fixation results represent only those trials that contained a clear sequence of left-to-right eye movements, interpretable in terms of the task. This pruning resulted in approximately 1% of the trials. Mean RT and inter-response interval for each subject (IRI) were computed only for stimuli correctly responded to.

Three fixation measures were computed: Eye-Hand Span (EHS), the time from the initial fixation on a stimulus till its response, Dwell, the duration of fixation on a stimulus, and Release-Hand Span (RHS), the time from when the eyes left stimulus N (presumably to fixate stimulus N+1) to when the response to stimulus N was made. For the first stimulus (S1), EHS is equivalent to RT. Overlap in processing adjacent stimuli means that EHS may reflect concurrent processing for multiple stimuli as well as scheduling strategies. Logically, EHS should be the sum of Dwell and RHS. In practice, each measure was computed independently from the eye movement samples

and manual responses. As a result, the sum of RHS and Dwell only approximates EHS. Dwell and RHS relate directly to the overlap in processing adjacent stimuli, since processing on stimulus N is still in progress while the eyes are fixated on N+1.

Analysis of Pure Sequences: We present first the analysis of the pure sequences, EEEEE and HHHHH, to assess the overall effects of difficulty. Means across subjects for manual responses (RT1/IRI) and Dwell are shown in Figure 2. Analysis of variance on mean RT1/IRI shows significant effects of position ($F[4,40] = 77, p < .001$) and difficulty ($F[1,10] = 18.3, p < .01$), but no position by difficulty interaction ($F[4,40] < 1$). Analysis of variance on mean Dwell for S1 through S4 found a trend in the effect of position ($F[3,30] = 2.1, p < .13$), a small but significant effect of difficulty ($F[1,10] = 5.1, p < .05$), but no interaction ($F[3,30] = 1.1, p > .3$).

The RT1 elevation and reduction in IRI for S5 ("last-item" effect) are typical findings. The pronounced increase in IRI from S2 to S4 has also been observed though not consistently. Mean Dwell time is relatively flat across position, though with a small elevation on S1. Since trial timing began with the onset of the sequence, Dwell and RT1 will be slightly elevated owing to the time required to recognize the display change and begin the task. We have previously observed that in the steady-state portion of sequence execution, generally from S3 to penultimate item, IRI and Dwell are approximately equal. This equivalence is also present here at S3 and S4.

Figure 3 shows the pattern of Eye-Hand Span (EHS) and Release-Hand Span (RHS) as a function of stimulus position for Hard and Easy sequences. Again, the pattern of EHS and RHS over items is very similar to that previously observed (see e.g., S.-C. Wu, & Remington, R. W., 2004). An analysis of variance on mean RHS showed a significant effect of position ($F[3,30] = 33.6, p < .001$), but no main effect of difficulty ($F[1,10] = 1.3, p > .20$) and no interaction between the two ($F[3,30] = 1.4, p > .25$). For EHS, the analysis of variance showed significant main effects of position ($F[3,30] = 31.3, p < .001$) and difficulty ($F[1,10] = 10.3, p < .01$), but no interaction ($F[3,30] < 1$). In the critical test, an analysis of variance of the pooled RHS of S3 and S4 failed to find a significant difficulty effect ($F[1,10] = 1.4, p > .25$).

Nonetheless, RHS shows an increasing effect of difficulty, reaching 100 ms at S4. This effect is due to 2 of the 11 subjects. For one subject, RHS in the Hard condition at those positions was 900 ms or more, far larger than the previous RHS, or any values previously seen for any other subject. It could indicate another strategic deferral of responses to later items in the sequence. For another subject, RHS in the

easy condition was negative, indicating that the response was made prior to the eye movement. Eliminating these scores and recomputing the anova confirmed that there was no effect of difficulty at these positions ($F[1,10] < 1$). The recomputed means for the Easy and Hard conditions at these positions were 225 ms and 235 ms, respectively. This strongly supports the idea that RHS is unaffected by difficulty and suggests that changes in processing after the eyes move is more a function of strategy than resource conflicts.

Analysis of Mixed-Difficulty Lists

To examine the effects of local and global context mixed lists were analysed as a function of item difficulty and difficulty of the context. Analysis of variance on mean IRI from S2 through S5 for each subject (omitting RT1) showed a significant effect of item difficulty ($F[1,10] = 41, p < .001$), but no effect of context (Hard, Easy) nor of the interaction of difficulty and context ($F < 1$ in both cases). A similar pattern was observed for Dwell: a significant effect of item difficulty ($F[1,10] = 26, p < .001$), no effect of context ($F < 1$), with a slight trend toward an interaction ($F[1,10] = 2.1, p < .2$). For RHS, none of the effects were significant. Again, RHS appears to be constant across both item difficulty and context. This analysis shows that the "global" context had no effect on IRI, Dwell, or RHS. These values showed none of the contextual assimilation effects that would be expected if subjects were using global estimates of processing time to schedule eye movements and manual responses.

The "local" context was examined by analysing the effects of preceding and subsequent stimuli. Effects of the previous item (N-1) on the Dwell and IRI of the current item (N) would disclose any push back effects arising from incomplete modulation of Dwell on N. For IRI, analysis of variance on the pooled values for S3 and S4 found, once again, a significant effect of difficulty of the current stimulus ($F[1,10] = 72, p < .001$), a non-significant trend in the effect of previous item difficulty ($F[1,10] = 2.6, p < .15$), and no interaction of difficulty and previous stimulus difficulty ($F < 1$). IRI was approximately 20 ms faster following an easy than a hard stimulus. For Dwell time there was a significant effect of difficulty ($F[1,10] = 25, p < .001$) but no effect of either previous item difficulty nor an interaction of current difficulty with previous difficulty ($F < 1$ in both cases). For RHS there were no significant effects ($F = 1$ for difficulty, $F < 1$ for the interaction), though there was a slight trend to shorter RHS when the previous items was easy ($F[1,10] = 2.2, p < .17$). This analysis is further evidence that difficulty even in mixed lists is almost fully

absorbed into the Dwell time so that when the eyes move a constant response-related act is all the processing that remains.

Likewise, analysis of variance showed no effect of the difficulty of N+1 on RHS, IRI, or Dwell for stimulus N ($F < 1$ in all cases). For IRI, there was a significant interaction of difficulty of the current item with difficulty of the subsequent item ($F[1,10] = 9.9, p < .01$). When the current item was Hard mean IRI was 674 ms and 710 ms for subsequent Easy and Hard stimulus, respectively. However, the interaction results from a small crossover effect: for easy items mean IRI was slower if the next item was easy than hard, 645 ms and 625 ms, respectively. Neither comparison was significant on post-hoc analysis.

General Discussion

Judgments of the number of items present (numerosity) are known to be affected by the compatibility with the number values represented. In keeping with this, incompatible numbers (saying "three" to three twos) produced substantial increases in inter-response intervals and dwell time, as well as increases in release-hand span, the time from when the eyes moved to the stimulus N+1 till when the response to N was made. Increases in RHS were confined to early stimulus position and, thus, potential inflated by voluntary strategies that defer initial responses. If subjects chose to fixate one or two subsequent items prior to making the first response, then the RHS to those items would be elevated whether or not there was significant processing during RHS. We have previously shown that by the third stimulus (S3) this strategy has transitioned into a steady-state phase of sequence execution characterized by regular saccades and manual responses (Remington et al). Analysis of RHS during this steady state portion shows no effect of difficulty. Interpreted narrowly, this indicates that our difficulty manipulation, the compatibility of the numbers with their numerosity, affected processing that took place during the eye fixation.

Interpreted more broadly, however, our results have important implications for the timing of saccadic eye movements across a range of sequence tasks. Our findings show that the difficulty of an individual item was fully absorbed into the fixation duration, so that once the eyes move there was a constant brief processing remaining, regardless of the difficulty of the item or context in which it occurred. This is strong support for active processing accounts and argues against models of global difficulty estimation. Because saccade programming takes time, it is quite likely that

preparation for its execution is based on an estimate of the completion of ongoing processing. Our results suggest that this estimation is based on the current state of processing, not that of previous (or expectations for subsequent) items.

An important caveat is that our results were obtained in a task in which people responded to each item in series. In that respect the task is like a simplified typing task. The overt response could have generated a focus treating each item in isolation, limiting overlap. Caution is required in generalizing to a task such as reading where each word is processed in turn, but where the goal is to comprehend the meaning given by the combination of words, not one word alone. For example, our data exhibit no "push back" effect of item difficulty on the dwell time of the subsequent item, while such effects have been reported for reading. Push back may reflect the integration of word meaning over lexical, syntactic, and semantic levels, some of which may occur only once the next word is fixated (i.e. during the RHS). That is, that the eyes maintain dwell until the relationship of the word to its preceding words and the overall meaning is understood. The problem with reading, however, is not just the lack of a clear marker of when the crucial central processes are completed, but a poor understanding of the rather complex process of comprehension.

There is still an issue as to whether the failure to find effects of difficulty on RHS is specific to the difficulty manipulation. As noted in the introduction, we (Wu, Remington, Pashler, 2004a, 2004b) reported an experiment in which stimulus-response compatibility was varied in separate sessions. Speeded forced choice key presses were made to the numbers, 1, 2, 3, or 4. In the compatible condition these numbers were assigned to the index, middle, ring, and little finger of the right hand, respectively. In the incompatible conditions the responses were scrambled. Unlike the present experiment, the effects of compatibility on RTI were more than double the effects on IRI. Still, stimulus response compatibility had no significant effect on RHS at positions 3 and 4, despite substantial effects on Dwell and IRI. The same pattern has been observed for luminance. We conclude from this that when people naturally execute a sequence, overlap is restricted to the response execution stage of one stimulus with the perceptual processing of the next.

Though the outcome provides qualitative support for the hypothesis that RS is largely complete prior to the saccade, we have simulated performance on the sequence to determine whether central bottleneck assumptions can provide good quantitative fits to the observed data. The task was simulated by estimating the stage

times for SE, RS, and RE, as well as central processing times for the saccade. We have previously estimated times for the strategic deferral of RT1 (Remington et al submitted) as consisting of start-up costs (60 ms) affecting both Dwell and RT1 plus a cost (100 ms) for the first time manual responses had to be coordinated with saccades. That earlier work also provided the estimate of saccade central processing (100 ms) and estimates for SE (200 ms) and RE (200 ms). To estimate RS for this task we used the model prediction that on the final item of the sequence $EHS = SE + RS + RE$, since there is no additional overhead for the saccade. RS was thus estimated by subtracting estimates of SE and RE from the final EHS separately for the hard and easy pure sequences. This also yielded the difficulty effect. Figure 4 plots predicted RT/IRI and Dwell as a function of position. Qualitatively the patterns are close to those observed. Quantitatively, the correlation of predicted and observed is quite high for RT1/IRI ($r = .97$) and EHS ($r = .98$), but only moderately high for Dwell ($r = .85$). The high correlations indicate that a simple stage model can accurately capture the patterns in the data.

Conclusions

The data are well fit by a central bottleneck model whose principal scheduling assumption is that RS is virtually complete by the time the saccade is made. The data show little if any effect of stimulus-response compatibility on processing that occurs after the saccade is made. It is possible to maintain that in fact saccade central processing is done before RS, just that the initiation of the saccade occurs after RS. It is difficult to see why that should be, but we leave that question to future research. At present, we can say with confidence that with no other pressure subjects will choose to move their eyes to the next stimulus only after determining the response to the current one. Whether this can be generalized to reading or other tasks that require the integration of items in a sequence remains to be determined.

References

- Hooge, L. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics*, 58(7), 969-976.
- Inhoff, A. W., & Gordon, A. M. (1997). Eye movements and eye-hand coordination during typing. *Current Directions in Psychological Science*, 6(6), 153-157.
- Inhoff, A. W., & Jian, W. (1992). Encoding of Text, Manual Movement Planning, and Eye Hand Coordination during Copytyping. *Journal of Experimental Psychology-Human Perception and Performance*, 18(2), 437-448.
- Inhoff, A. W., Rosenbaum, D. A., Gordon, A. M., & Campbell, J. A. (1984). Stimulus-Response Compatibility and Motor Programming of Manual Response Sequences. *Journal of Experimental Psychology-Human Perception and Performance*, 10(5), 724-733.
- John, B. E. (1996). TYPIST: A theory of performance in skilled typing. *Human-Computer Interaction*, 11(4), 321-355.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25-26), 3559-3565.
- Land, M. F., & LEE, D. N. (1994). Where We Look When We Steer. *Nature*, 369(6483), 742-744.
- Land, M. F., & Mcleod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience*, 3(12), 1340-1345.
- Legge, G. E., Hooven, T. A., Klitz, T. S., Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Research*, 42(18), 2219-2234.
- Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr Chips: An ideal-observer model of reading. *Psychological Review*, 104(3), 524-553.
- Pashler, H. (1984). Processing Stages in Overlapping Tasks - Evidence for a Central Bottleneck. *Journal of Experimental Psychology-Human Perception and Performance*, 10(3), 358-377.
- Pashler, H. (1994). Overlapping Mental Operations in Serial Performance with Preview. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 47(1), 161-191.
- Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research*, 139(3), 266-277.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.
- Rayner, K., & Pollatsek, A. (1981). Eye-Movement Control during Reading - Evidence for Direct Control. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 33(NOV), 351-373.
- Reichle, E. D., & Laurent, P. A. (2006). Using reinforcement learning to understand the emergence of "intelligent" eye-movement behavior during reading. *Psychological Review*, 113(2), 390-408.

- Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, 7(1), 4-22.
- Remington, R. W., Lewis, R., & Wu, S.-C. (2006). Scheduling mental operations in a multiple-response sequence: Modeling the effects of a strategy to minimize variance in the timing of saccades. In D. Fum, F. del Missier, & A. Stocco (Ed.), *Proceedings of the 7th International Conference on Cognitive Modeling, Trieste, Italy. April 2-6, 2006*.
- Salthouse, T. A. (1986). Perceptual, Cognitive, and Motoric Aspects of Transcription Typing. *Psychological Bulletin*, 99(3), 303-319.
- Salthouse, T. A., & Ellis, C. L. (1980). Determinants of Eye-Fixation Duration. *American Journal of Psychology*, 93(2), 207-234.
- Salthouse, T. A., Ellis, C. L., Diener, D. C., & Somberg, B. L. (1981). Stimulus-Processing during Eye Fixations. *Journal of Experimental Psychology-Human Perception and Performance*, 7(3), 611-623.
- Sanders, A. F., & van Duren, L. L. (1998). Stimulus control of visual fixation duration in a single saccade paradigm. *Acta Psychologica*, 99(2), 163-176.
- Sternberg, S. (1969). Discovery of Processing Stages - Extensions of Donders Method. *Acta Psychologica*, 30, 276-+.
- Vickers, J. N. (1992). Gaze Control in Putting. *Perception*, 21(1), 117-132.
- Welford, A. T. (1952). The 'psychological refractory period' and the timing of high-speed performance -- a review and a theory. *British Journal of Psychology*, 43, 2-19.
- Wu, S.-C., & Remington, R. W. (2004). Coordination of component mental operations in a multiple-response task. In S. N. Spencer (Ed.), *Proceedings of the Eye Tracking Research and Applications Symposium 2004. New York: ACM SIGGRAPH*.
- Wu, S.-C., Remington, R. W., & Pashler, H. (2004). Coordination of component mental operations in sequences of discrete responses. In K. Forbus, D. Gentner, & T. Regier (Ed.), *Proceeding of the Twenty-Sixth Annual Conference of the Cognitive Science Society (pp. 1476-1481). Mahwah, NJ: LEA*.
- Wu, S.-C., Remington, R.W., & Pashler, H. (2007). Scheduling of Eye Movements and Manual Responses in Performing a Sequence of Choice Responses: Empirical Data and Model. In R. L. Lewis, T.A. Polk, & J.E. Laird (Ed.), *Proceedings of the 8th International Conference on Cognitive Modeling. Ann Arbor, MI July 26-29, 2007*.
- Wu, S.-C., Remington, R.W., Lewis, R. (2007). Modeling the scheduling of eye movements and manual responses in performing a sequence of discrete trials. In R. Sun (Ed.), *Proceedings of the 28th Annual Cognitive Society, Vancouver, BC, Canada, July 26-29*.

Figure Captions

Figure 1:

The four panels depict different schedules for coordinating eye movements with ongoing task processing. In all panels SE = Stimulus Encoding, RS = Response Selection, EM = Eye Movement (central processor portion), and RE = Response Execution. The dependent measures of Dwell (eye fixation duration) and RHS (Release-Hand Span) are shown for all; IRI (inter-response interval) and EHS (eye-hand span) are shown in the first panel.

(A): A schedule in which RS for task processing is completed before the central processing of the eye movement. RHS consists only of the RE stage and only RE and SE are overlapped.

(B): A schedule in which the eye movement is made prior to RS. RHS is longer due to the RS stage, while Dwell also includes a portion of RS as the EM stage must queue.

(C) & (D): These show cases in which the processing is interrupted to generate an eye movement timed to minimize delays in central stages. Completion of the RS stage is estimated from either ongoing processing or immediate context, and the eye movement timed so that RS on N and SE on N+1 complete at the same time. Panel (C) shows a case where both N and N+1 have the same RS duration. Panel (D) shows a case where RS on N is shorter than N+1. In (D) the effects of context cause an underestimate of the needed Dwell on N+1, lengthening the RHS for that stimulus.

Figure 2:

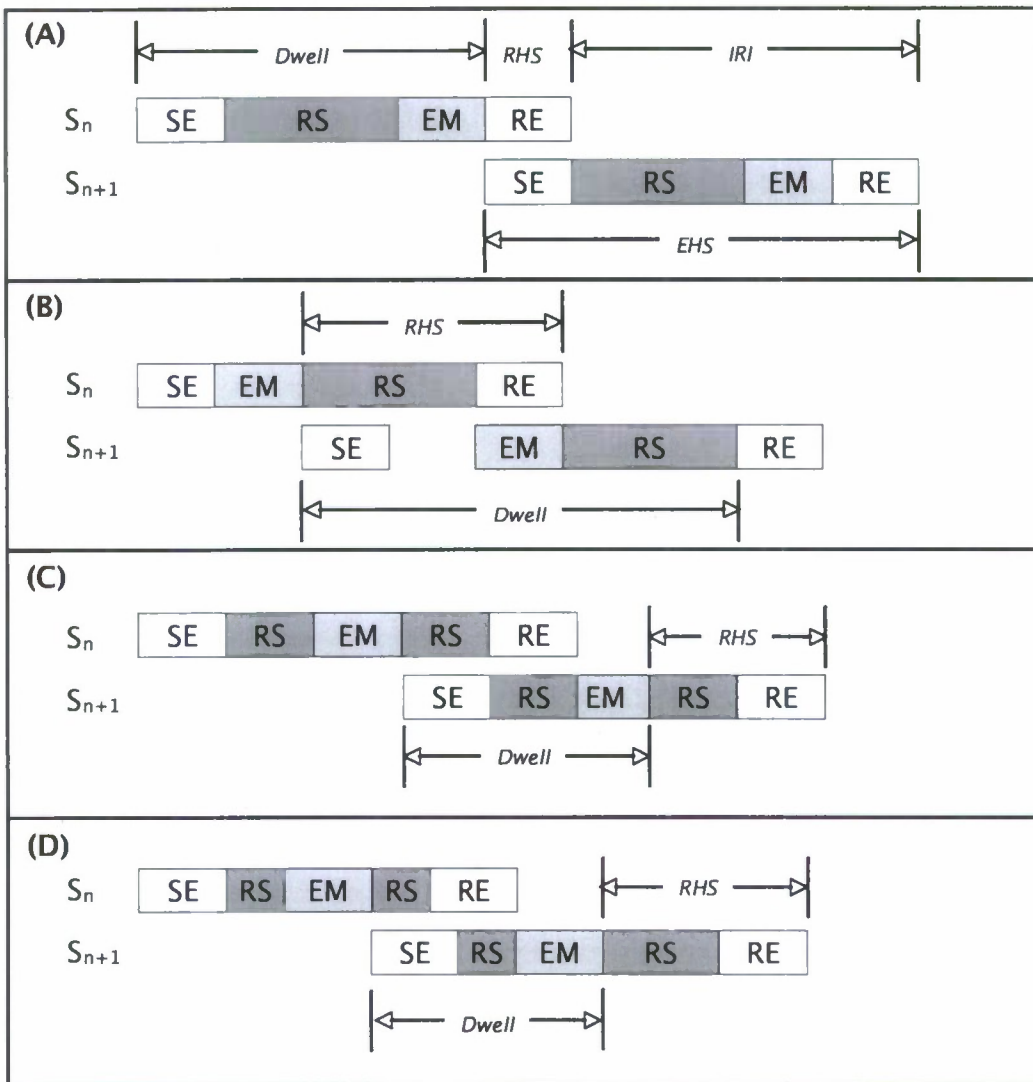
Manual responses (RT1/IRI) and Dwell for the present experiment as a function of stimulus position.

Figure 3:

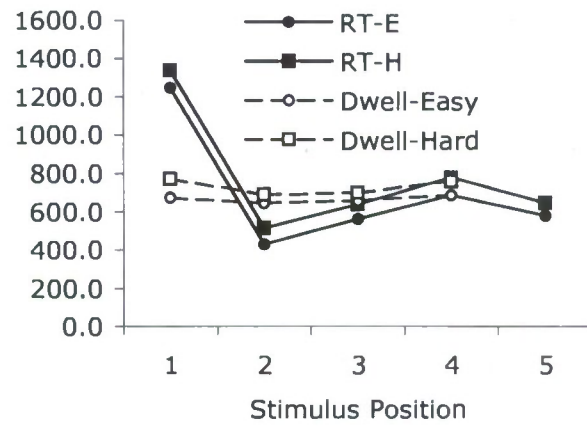
EHS and RHS for the present experiment.

Figure 4:

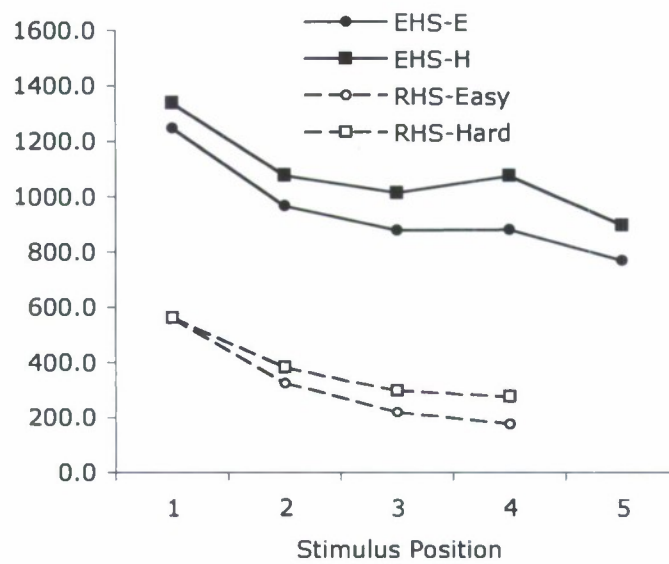
Predicted RT1/IRI and Dwell from an RS-S model adapted from Figure 1.



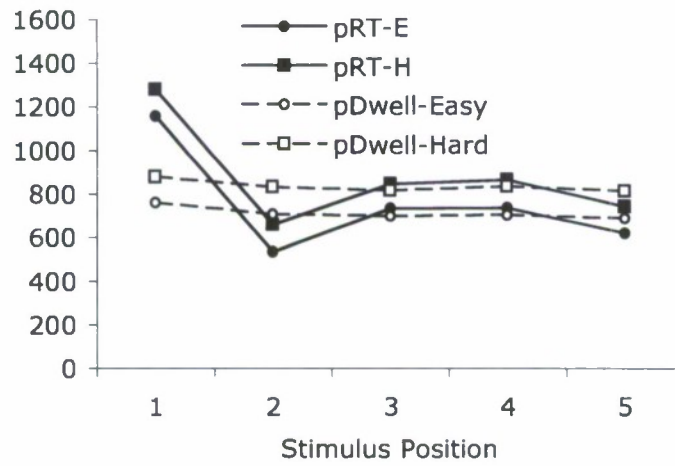
RT and Dwell for Pure Sequences



EHS and RHS for Pure Sequences



RT and Dwell for Pure Sequences



AFOSR Finding Happiness

Research Question:

An operator's effectiveness at monitoring multiple signals increases with experience. Using eye-tracking during an experimental task, we will describe how this skill develops. Ultimately, we seek to answer the question: how can we speed the rate at which an operator adapts his monitoring behavior to optimally suit his environment?

Experiment Overview:

The subject's task is simply to report the appearance of a happy face ☺ on the computer monitor. What will make this task non-trivial, however, is the fact it will be unknown to the subject when and where the face will appear, and that all but a small portion of the screen (the area the subject chooses to monitor) will be obscured until the subject chooses a new area to monitor. This experiment will allow for the characterization of the learning process subjects go through to adapt their monitoring behavior to the dynamics of a stochastically operating system.

Application Specifications:

The experiment application will first register the subject and create a data file to be written to periodically throughout the procedure. Existing programs in this lab can be used as a model for this stage.

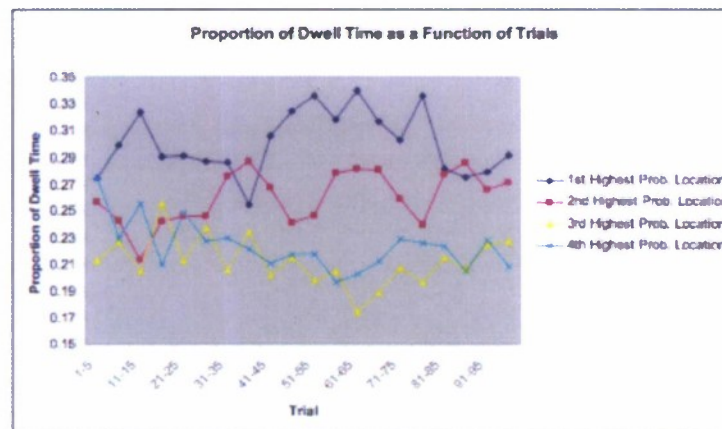
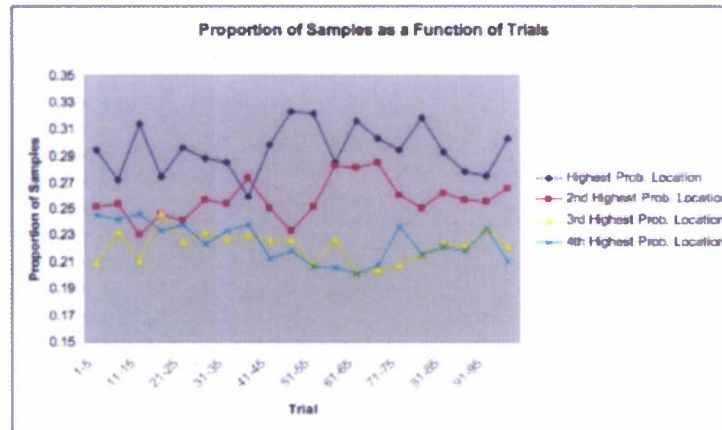
In the task proper, subjects will confront a screen as depicted here:

Finding Happiness Data Review

9/4/2007

Summary

Subjects demonstrate learning by adjusting their sampling behavior as trials progress. Subjects visually sample high-probability target areas more often than low-probability target areas. This leads to greater total gaze dwell times in high probability target areas.



Optimal Dwell Duration Dependent on Saccade Duration

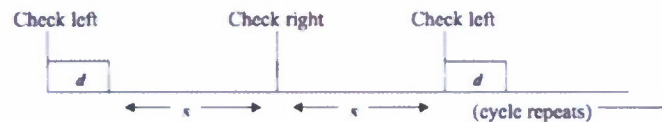
Based on a discussion Hal and I had recently, I've developed below a mathematical treatment of a simplified version of the Finding Happiness experiment. I believe it will prove useful in two ways:

- 1) It verifies and quantifies some of the intuitions we've discussed about what subjects "should" do after learning the probabilities of different target locations.
- 2) It allows for further hypothesis generation. One implication of the model below that I don't think we've discussed is that the benefit of adjusting one's dwell time in accordance with the target location probabilities grows as the time to saccade between target locations increases. (More on that point later.)

First, some assumptions and definitions:

- This model assumes only TWO possible target locations. We can refer to them as "left" and "right."
- p is the probability that the target will appear in the left location. $(1-p)$ is the probability the target will appear in the right location.
- s is the saccade time. This is the amount of time spent moving the gaze from one location to the other. No detection of the target is possible in this time. It is assumed s is symmetrical (e.g. the time to move left to right equals the time to move right to left).
- d is the additional left dwell time. This is the additional time spent gazing at the left location prior to saccading away to the right. For example, if $d = 0$, then the time spent examining left and right locations is equal. If $d = 100$, then 100 milliseconds will be spent waiting for the target to appear in the left location before saccading to the right location.
- Detection of the target is assumed to be instantaneous (takes 0 milliseconds) as is the generation of the response. (Variables may be substituted for these values later, but the effect would be to simply add a constant to all predicted values from this model.)

A diagram best describes how we will come to calculate expected RTs from this model:



Note again that the "check" states are instantaneous – if the target is present it is immediately detected. If not, the model moves immediately to the next state (a d or s period).